

```
QY 362 TTAGCAATTAATGACAACTATGAGCTTTAAAGTCATCAGCATGATGACAGAGAG 421
DB 506 TTAGCAATTAATGACAACTATGAGCTTTAAAGTCATCAGCATGATGACAGAGAG 565
QY 422 GAGTCCATTATGAGTATCCGAGAGCTTCTCTCGTGAAGGGTTGAAACATGCCAATA 481
DB 566 GAGTCCATTATGAGTATCCGAGAGCTTCTCTCGTGAAGGGTTGAAACATGCCAATA 625
QY 482 TTGTGCTCTGACATGACATATATCACAACCAAGACACTGACATTCGTTTTGAATACA 541
DB 626 TTGTGCTCTGACATGACATATATCACAACCAAGACACTGACATTCGTTTTGAATACA 685
QY 542 TGACACAGACCTGCGCCAGTATATGTCTCAGCATCCAGAGGGCTTCATCTCATATG 601
DB 686 TGACACAGACCTGCGCCAGTATATGTCTCAGCATCCAGAGGGCTTCATCTCATATG 745
QY 602 TCAGACTTTTCAATGTTTCAACTTTTGGGGGCTGGGGTATCCACCAACCAAGGTTG 661
DB 746 TCAGACTTTTCAATGTTTCAACTTTTGGGGGCTGGGGTATCCACCAACCAAGGTTG 805
QY 662 TTACACAGGACCTGAAACCTCAGAACTTACTCATCAGTCACTGGAGAGCTCAAACTGG 721
DB 806 TTACACAGGACCTGAAACCTCAGAACTTACTCATCAGTCACTGGAGAGCTCAAACTGG 865
QY 722 CTGATTTTGGTCTTGGCCGGGCGCAAGTCCATTCACGACGACATCTTTCAGAAAGTCG 781
DB 866 CTGATTTTGGTCTTGGCCGGGCGCAAGTCCATTCACGACGACATCTTTCAGAAAGTCG 925
QY 782 TGACCTCTGTTATCCGGGCCCCCTGATGCTTGTGGAGGACATGAAATATCTCTGAGC 841
DB 926 TGACCTCTGTTATCCGGGCCCCCTGATGCTTGTGGAGGACATGAAATATCTCTGAGC 985
QY 842 TGACATATGAGGCTGAGGCTGATCTTTATGAAATGTTCCAGGGTCAACCTTTGTTTC 901
DB 986 TGACATATGAGGCTGAGGCTGATCTTTATGAAATGTTCCAGGGTCAACCTTTGTTTC 1045
QY 902 CTGGGGTTTTCACATCTTTTGAACAGCTGAGAGAAATCTGGAGGGTCTGGAGGTCCTTA 961
DB 1046 CTGGGGTTTTCACATCTTTTGAACAGCTGAGAGAAATCTGGAGGGTCTGGAGGTCCTTA 1105
QY 962 CAGAGGATATCTGGGCGGAGGTCTCCAGCTACTTAACATTCAGAAATGCT 1015
DB 1106 CAGAGGATATCTGGGCGGAGGTCTCCAGCTACTTAACATTCAGAAATGCT 1159

RESULT 12
ADM16422
ID ADM16422 standard; DNA; 1828 BP.
XX
AC ADM16422;
XX
XX
DT 17-JUN-2004 (first entry)
XX
DE DNA encoding human kinase protein, seq id 1.
XX
XX Cytostatic; cardiant; neuroprotective; nootropic; antiinfertility;
XX vulnery; antidiabetic; kinase; cancer; heart disease;
XX Alzheimer's disease; infertility; wound; diabetes; neurological disease;
XX gene; de.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX CDS 87..1124
XX FT /tag= a
XX FT /product= "kinase"
XX
XX WO2004024913-A1.
XX
XX 25-MAR-2004.
XX
XX 10-SEP-2003; 2003WO-JP011552.
```

```
XX
PR 10-SEP-2003; 2002JP-00264345.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX (ZOE-) ZOEGENE CORP.
XX
PI Kondo J, Kawai K, Miyama N, Nakajima M, Isogai T, Sugiyama T;
PI Wakamatsu A, Irie R, Ishii S;
XX
XX MPI; 2004-270042/25.
XX
XX P-PSDB; ADM16427.
XX
XX Proteins of human origin having kinase activity, useful for prevention
XX PT and treatment of kinase-associated diseases including cancer, heart
XX PT disease and Alzheimer's disease.
XX
XX Claim 4; SEQ ID NO 1; 105bp; Japanese.
XX
XX The invention relates to four proteins of human origin (I) having kinase
XX CC activity, and to proteins derived from these by addition, deletion and/or
XX CC substitution of one or more amino acid residues, and having similar
XX CC activity. The proteins and other aspects of the invention are useful for
XX CC the prevention and treatment of kinase-associated diseases including
XX CC cancer, heart disease, Alzheimer's disease, infertility, wounds, diabetes
XX CC and neurological diseases. The current sequence represents a human DNA
XX CC encoding a protein having kinase activity.
XX
XX Sequence 1828 BP; 496 A; 398 C; 450 G; 484 T; 0 U; 0 Other;
XX
XX Query Match 66.7%; Score 873; DB 12; Length 1828;
XX Best Local Similarity 93.8%; Pred. No. 9.8e-269;
XX Matches 946; Conservative 0; Mismatches 0; Indels 63; Gaps 1;
XX
QY 1 ATGGGTCAGAGCTGTGTGCAAAAGTGTACAGCTGTAGTACAGTGTCTACATTGTTCA 60
DB 87 ATGGGTCAGAGCTGTGTGCAAAAGTGTACAGCTGTAGTACAGTGTCTACATTGTTCA 146
QY 61 GAGGAGGCGGAGGACACACAGCTGTGGAGAGTCAAGCTGTAGACCAAGAGGCTGCTTC 120
DB 147 GAGGAGGCGGAGGACACACAGCTGTGGAGAGTCAAGCTGTAGACCAAGAGGCTGCTTC 206
QY 121 AAGCTTAACAGACTTAATAAAGATCATGTTCCATGACTTCAATTCACCCAGGGGACTT 180
DB 207 AAGCTTAACAGACTTAATAAAGATCATGTTCCATGACTTCAATTCACCCAGGGGACTT 266
QY 181 CAAGTGCCTGCGCCAGAAAGTTCAAGAGTAAAGGCGCAGGATTAAGTATGTTT 240
DB 267 CAAGTGCCTGCGCCAGAAAGTTCAAGAGTAAAGGCGCAGGATTAAGTATGTTT 326
QY 241 CAGAAAGAGGATCTGAGCGAGGGTTTTCAGTGGAGAAAGGCTCTTTTGGGGGAGCC 300
DB 327 CAGAAAGAGGATCTGAGCGAGGGTTTTCAGTGGAGAAAGGCTCTTTTGGGGGAGCC 386
QY 301 TCATCTTACTTGAATCTTGAAGAAAGTGGGTGAAGGCTCTTAAGCCAGTTTAAAGGGG 360
DB 387 TCATCTTACTTGAATCTTGAAGAAAGTGGGTGAAGGCTCTTAAGCCAGTTTAAAGGGG 446
QY 361 ATTAGAGAAATTAATGAGCAACTAGTGGCTTTAAAGTCAATGAGATGATGAGAGAA 420
DB 447 ATTAGAGAAATTAATGAGCAACTAGTGGCTTTAAAGTCAATGAGATGATGAGAGAA 506
QY 421 GGAGTCCCAATTTACAGCTATCCGAGAAAGCTTCTCTTGAAGGGTTTGAAGATGCAAT 480
DB 507 GGAGTCCCAATTTACAGCTATCCGAGAAAGCTTCTCTTGAAGGGTTTGAAGATGCAAT 566
QY 481 ATTGTGCTCTGCAATGACATTAATCCACCAAGAGACATGACATTCGTTTTGAATGC 540
DB 567 ATTGTGCTCTGCAATGACATTAATCCACCAAGAGACATGACATTCGTTTTGAATGC 626
QY 541 ATGACACAGACCTGGCCAGTATATGTTCTCAGCATCCAGAGGGCTTCATCTCATTAAT 600
DB 627 ATG----- 629
```

QY 601 GTCAGACTTTTCACTTTTCACTTTTGGGGGCTGGCTGATCCAGCAACCAACGCTT 660
 DB 630 -----CTTTTCAATGTTTCAACTTTTGGGGGCTGGCTGATCCAGCAACCAACGCTT 683
 QY 661 CTTTCAAGGAGACTGAAACCTTCAGACTTACTCATGATCACTGGGAGAGCTCAAACTG 720
 DB 684 CTTTCAAGGAGACTGAAACCTTCAGACTTACTCATGATCACTGGGAGAGCTCAAACTG 743
 QY 721 GCTGATTTTGTGCTTGGCCCGGGGCAAGTCATTCAGCCAGCATATCTTTCAGAAATG 780
 DB 744 GCTGATTTTGTGCTTGGCCCGGGGCAAGTCATTCAGCCAGCATATCTTTCAGAAATG 803
 QY 781 GTGACCTCTGTGATCCGGCCCTGTGATGCTTGTGAGGCCATGAATATTCCTGTAG 840
 DB 804 GTGACCTCTGTGATCCGGCCCTGTGATGCTTGTGAGGCCATGAATATTCCTGTAG 863
 QY 841 CTTGACATATGAGGATGAGCTGATCTTATTTGAATGTTTCAGAGGTCAACTTTGTTT 900
 DB 864 CTTGACATATGAGGATGAGCTGATCTTATTTGAATGTTTCAGAGGTCAACTTTGTTT 923
 QY 901 CTTGGGGTTTCCAAATCTTGTGAACAGCTGGAGAAAATCTGGAGGTGCTGGAGTCCCT 960
 DB 924 CTTGGGGTTTCCAAATCTTGTGAACAGCTGGAGAAAATCTGGAGGTGCTGGAGTCCCT 983
 QY 961 ACAGAGATATCTTGGCCGGAGTCTCCAGCTACTCAATCAATCAG 1009
 DB 984 ACAGAGATATCTTGGCCGGAGTCTCCAGCTACTCAATCAATCAG 1032

RESULT 13

AB277160
 ID AB277160 standard, cDNA, 888 BP.

AC AB277160;

DT 07-MAY-2003 (first entry)

XX Human protein kinase encoding cDNA SEQ ID NO:69.

XX Human; protein kinase; enzyme; antidiabetic; antiinflammatory;
 XX antidiabetic; antiparkinsonian; antidiabetic; cardiant; cyostatic;
 XX immunosuppressive; vulnereary; gene therapy; COPD; asthma; migraine;
 XX chronic obstructive pulmonary disease; non-insulin dependent diabetes;
 XX Parkinson's disease; myocardial infarction; inflammatory bowel disease;
 XX autoimmune disorder; allograft rejection; graft versus host disease;
 XX cancer; leukemia; wound granulation; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..888

XX /tag= a

XX /partial

XX /product= "protein kinase"

XX /note= "no start or stop codons given"

XX WO2003000901-A2.

XX 03-JAN-2003.

XX 24-JUN-2002; 2002WC-IB002358.

XX 26-JUN-2001; 2001US-0301098P.

XX 06-NOV-2001; 2001US-0332870P.

XX (DECO-) DECODE GENETICS EHF.

XX Martinez RAM, Sigurdson GT;

XX WPI; 2003-201429/19.

XX P-PSDB; ABP96082.

XX New protein kinase genes and polypeptides, useful for diagnosing diseases

PT associated with a protein kinase, or in gene therapy for treating e.g.
 PT Parkinson's disease, migraine, myocardial infarction, allograft rejection
 PT or cancers.

XX Claim 1; Page 85; 258bp; English.

XX AB277160 to AB277165 encode the human protein kinases given in ABP96048
 CC to ABP96087. The protein kinases have antidiabetic, antiinflammatory,
 CC antidiabetic, antiparkinsonian, antidiabetic, cardiant, cyostatic,
 CC immunosuppressive and vulnereary activities, and can be used in gene
 CC therapy. A protein kinase therapeutic agent from the present invention,
 CC particularly a protein kinase gene agonist or antagonist, can be used for
 CC treating a disease or condition associated with a protein kinase in an
 CC individual. These diseases include chronic obstructive pulmonary diseases
 CC (COPD), asthma, non-insulin dependent diabetes, Parkinson's disease,
 CC migraine, myocardial infarction, inflammatory bowel disease, autoimmune
 CC disorders (e.g. allograft rejection or graft vs. host disease), cancers
 CC (e.g. leukemia) or wound granulation

SQ Sequence 888 BP; 214 A; 234 C; 207 G; 233 T; 0 U; 0 Other;

Query Match 62.8%; Score 821; DB 10; Length 888;

Best Local Similarity 95.7%; Pred. No. 3.2e-252;

Matches 888; Conservative 0; Mismatches 30; Indels 9; Gaps 1;

QY 298 GCCTCATCTTACTTGAACCTTGGAGAGCTGGGTGAAGCTCTTATGCGACATTTTACAG 357
 DB 1 GCCTCATCTTACTTGAACCTTGGAGAGCTGGGTGAAGCTCTTATGCGACATTTTACAG 60
 QY 358 GGGATTAGAGAAATTAATGAGCACTAGTGGCTTTTAAAGTCATGATGATGATGATGATG 417
 DB 61 GGGATTAGAGAAATTAATGAGCACTAGTGGCTTTTAAAGTCATGATGATGATGATGATG 120
 QY 418 GAAAGATGCCATTTTACAGCTATCCGAGAGCTTCTCTGTAAGGTTGAAATATGCC 477
 DB 121 GAAAGATGCCATTTTACAGCTATCCGAGAGCTTCTCTGTAAGGTTGAAATATGCC 180
 QY 478 AATATTGGCTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
 DB 181 AATATTGGCTCTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 538 TACATGACACAGAGACTGGCCCAATATATGTCAGATCCAGAGGGCTTATCTCAT 597
 DB 241 TACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 598 AATGTCAGCTTTTCAATGTTTCAACTTTTGGCGGGCTGGCTATCATCCACCAACAC 657
 DB 301 AATGTCAGCTTTTCAATGTTTCAACTTTTGGCGGGCTGGCTATCATCCACCAACAC 360
 QY 658 GTTCTTCAAGAGGAGCTGAAACCTCAGAACTTACTCATGATGATGATGATGATGATGATGAT 717
 DB 361 GTTCTTCAAGAGGAGCTGAAACCTCAGAACTTACTCATGATGATGATGATGATGATGATGAT 420
 QY 718 CTGGCTGATTTTGTCTTGGCCCGGGCAAGTCATTTCCAGCCAGCATATCTTTCAGAA 777
 DB 421 CTGGCTGATTTTGTCTTGGCCCGGGCAAGTCATTTCCAGCCAGCATATCTTTCAGAA 480
 QY 778 GTGCTGACCTCTGATCCGGCCCTGTATGCTTTTGGGAGCCACTGAATATTCCTT 837
 DB 481 GTGCTGACCTCTGATCCGGCCCTGTATGCTTTTGGGAGCCACTGAATATTCCTT 540
 QY 838 GAGCTGACATATGAGGAGTGCAGAGCTGATCTTATTTGAATGTTTCAGGGTCAACTTTG 897
 DB 541 GAGCTGACATATGAGGAGTGCAGAGCTGATCTTATTTGAATGTTTCAGGGTCAACTTTG 600
 QY 898 TTTCTGAGGTTTCCAAATCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTTC 957
 DB 601 TTTCTGAGGTTTCCAAATCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTTC 660
 QY 958 CTTACAGAGATATCTTGGCCGGAGTCTTCAAGCTTAACTTAACTTAACTTAACTTAACTTAA 1017
 DB 661 CTTACAGAGATATCTTGGCCGGAGTCTTCAAGCTTAACTTAACTTAACTTAACTTAACTTAA 720

QY 1018 CCAGTCCCTACGCGCTCGAAGCCTTCATGTGTCTGGAACAGGCTGGGAGGGTCTCTGAA 1077
 Db 721 AGAACTCATATTTTCTTCTCC-----ACTTTTCCAGGCTGGGAGGGTCTCTGAA 771
 QY 1078 GCTGAAGACCTGCGCTCCAGATGCTPAAAGGCTTTCCAGAGACCGCGTCTCCGCCAG 1137
 Db 772 GCTGAAGACCTGCGCTCCAGATGCTPAAAGGCTTTCCAGAGACCGCGTCTCCGCCAG 831
 QY 1138 GAAGCATTGTCATATTTTACAGCGCCCTGCTCCATCTCAGCTGTACCACTTCTCT 1194
 Db 832 GAAGCATTGTCATATTTTACAGCGCCCTGCTCCATCTCAGCTGTACCACTTCTCT 888
 RESULT 14
 AAF44668
 AAF44668 standard; cDNA; 1077 BP.
 AC AAF44668;
 XX
 XX 27-MAR-2001 (first entry)
 DT
 XX Novel protein kinase cDNA, SEQ ID NO: 48.
 DE
 XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
 KM immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
 KM dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
 KM immune disorder; cardiovascular disease; neurodegenerative disease;
 KM cancer; autoimmune disorder; stroke; inflammatory bowel disease;
 KM inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200073469-A2.
 PD 07-DEC-2000.
 XX
 XX 26-MAY-2000; 2000MO-US014842.
 PF
 XX 28-MAY-1999; 99US-0136503P.
 PR
 XX (SUGEN-) SUGEN INC.
 PA
 XX Plowman GD, Martinez R, Whyte D, Suderanam S;
 PI WPI; 2001-032161/04.
 DR P-PSDB; AAB65641.
 DR
 XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and
 PT treating immune-related diseases and disorders, cardiovascular disease,
 PT neurodegenerative diseases and/or cancers.
 PT
 XX Example 1; Fig 2; 310pp; English.
 PS
 XX The present sequence encodes a novel protein kinase. The nucleic acids
 CC and the protein kinases they encode may be used in the treatment and
 CC diagnosis of diseases associated with inappropriate kinase expression
 CC such as immune-related diseases and disorders, cardiovascular disease,
 CC neurodegenerative diseases and/or cancers. The nucleic acids and
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC assays. The kinase polypeptides may be used as antigens in the production
 CC of antibodies of kinase expression and activity. Anti-kinase antibodies
 CC and kinase antagonists may also be used to down regulate kinase
 CC expression and activity. Diseases related to kinase expression and
 CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
 CC disorders, complications of organ transplantation, myocardial infarction,
 CC immune disorders, cardiomyopathies, strokes, renal failure, oxidative-
 CC stress related disorders, chronic inflammatory bowel disease, chronic
 CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
 CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
 CC disorders
 XX
 SQ Sequence 1077 BP; 269 A; 257 C; 248 G; 303 T; 0 U; 0 Other;

Query Match 53.7%; Score 702.2; DB 4; Length 1077;
 Best Local Similarity 99.6%; Pred. No. 4.9e-214;
 Matches 704; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 602 TCAGACTTTTCATGTTTCAACTTTTGGCGGGCTGGCGTACATCAACCAACAGTTTC 661
 Db 95 TCAGACTTTTCATGTTTCAACTTTTGGCGGGCTGGCGTACATCAACCAACAGTTTC 154
 QY 662 TTCACAGGAGACCTGAAACCTCAAGAACTTACTATCATCACTCTGGAGAGCTCAACAG 721
 Db 155 TTCACAGGAGACCTGAAACCTCAAGAACTTACTATCATCACTCTGGAGAGCTCAACAG 214
 QY 722 CTGATTTTGGTCTTGCGCGGAGCAAGTCCATTCGCCAGCAGATACCTTCAGAAAGTGG 781
 Db 215 CTGATTTTGGTCTTGCGCGGAGCAAGTCCATTCGCCAGCAGATACCTTCAGAAAGTGG 274
 QY 782 TGACCTCTGTGATACCGGCGCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTGTAGC 841
 Db 275 TGACCTCTGTGATACCGGCGCCCTGATGCTTTGCTGGAGCCACTGAATATTCCTGTAGC 334
 QY 842 TGGACATATGCGGTGCAAGCTGCATCTTATTTGAATGTTCCAGGGTCAACCTTGTTC 901
 Db 335 TGGACATATGCGGTGCAAGCTGCATCTTATTTGAATGTTCCAGGGTCAACCTTGTTC 394
 QY 902 CTGGGGTTTCCAAACATCCTTGAACAGCTGGAGAAATCTGGAGGTGCTGGAGTCCCTA 961
 Db 395 CTGGGGTTTCCAAACATCCTTGAACAGCTGGAGAAATCTGGAGGTGCTGGAGTCCCTA 454
 QY 962 CAGAGATACCTTGGCGGGAGTCTCCAACTACCTAATCAATCAGAAATGTTTCCAC 1021
 Db 455 CAGAGATACCTTGGCGGGAGTCTCCAACTACCTAATCAATCAGAAATGTTTCCAC 514
 QY 1022 TGCCTACGCTCTGAAAGCCTTCAATGTTTGTGGAACAGCTGGGCGAGGTTCTGAAAGCTG 1081
 Db 515 TGCCTACGCTCTGAAAGCCTTCAATGTTTGTGGAACAGCTGGGCGAGGTTCTGAAAGCTG 574
 QY 1082 AAGACCTGGCTCCAGATGCTPAAAGGCTTTCCAGAGACCGCGTCTCCGCCAGGAAG 1141
 Db 575 AAGACCTGGCTCCAGATGCTPAAAGGCTTTCCAGAGACCGCGTCTCCGCCAGGAAG 634
 QY 1142 CACTTGTTCATATTTTCAAGCGCCCTGCTCAGCTGTACCACTTCTGTAGAG 1201
 Db 635 CACTTGTTCATATTTTCAAGCGCCCTGCTCAGCTGTACCACTTCTGTAGAG 694
 QY 1202 AGCTTTGTTTACAGTTTCAAGAGTGAAGCTTAAAGCAGAAATGTGTGACCTTTGGCCT 1261
 Db 695 AGCTTTGTTTACAGTTTCAAGAGTGAAGCTTAAAGCAGAAATGTGTGACCTTTGGCCT 754
 QY 1262 CCTACCGAAGGTACCAAGCCAGGCTTAAAGAAATGCTGTGA 1308
 Db 755 CCTACCGAAGGTACCAAGCCAGGCTTAAAGAAATGCTGTGA 801
 RESULT 15
 AD129366
 ID AD129366 standard; cDNA; 1077 BP.
 XX
 XX AD129366;
 AC
 XX 22-APR-2004 (first entry)
 DT
 XX Human MARK3-associated cDNA #36.
 DE
 XX Human; ss; antisense gene therapy; MARK3;
 KM MAP/microtubule affinity-regulating kinase 3; cancer;
 KM Alzheimer's disease; neurodegenerative disorder;
 KM hyperproliferative disorder; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN US2003232771-A1.
 XX
 XX 18-DEC-2003.

XX 17-JUN-2002; 2002US-00174319.
XX 17-JUN-2002; 2002US-00174319.
XX (ISIS-) ISIS PHARM INC.
XX Ward DT, Freier SM, Dobie KW;
XX WPI; 2004-052188/05.
XX P-PSDB; ADI29248.
XX
XX New antisense compound targeted to a nucleic acid molecule encoding
XX microtubule-affinity-regulating kinases (MARK3), useful for modulating
XX expression of MARK3 or for treating cancer or Alzheimer's disease.
XX
XX Disclosure; Fig 2; 233pp; English.
XX
XX The invention relates to a compound comprising a sequence comprising 8-80
XX base pairs (bp) targeted to a nucleic acid encoding MARK3
XX (MAP/microtubule affinity-regulating kinase 3), that specifically
XX hybridizes with the nucleic acid encoding MARK3 and inhibits expression
XX of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a
XX composition comprising the compound and a carrier or diluent, inhibiting
XX the expression of MARK3 in cells or tissues, treating an animal having or
XX suspected of having a disease or condition associated with MARK3 and
XX screening for an antisense compound. The antisense oligonucleotide is
XX useful for preparing a composition for treating hyperproliferative
XX disorder, particularly cancer and neurodegenerative diseases e.g.
XX Alzheimer's disease. The present sequence is a MARK3 associated cDNA
XX included in the figures but not mentioned anywhere else in the
XX specification.
XX
XX Sequence 1077 BP; 269 A; 257 C; 248 G; 303 T; 0 U; 0 Other;

Query Match 53.7%; Score 702.2; DB 12; Length 1077;
Best Local Similarity 99.6%; Pred. No. 4,9e-214;
Matches 704; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 602 TCAGACTTTTCATGTTTCAACTTTTGCGGGGCTGGGCTATCCACACCAACAGCTTC 661
DB 95 TCCAGCTTTTCATGTTTCAACTTTTGCGGGGCTGGGCTATCCACACCAACAGCTTC 154
QY 662 TTCACAGGAGCTCGAAGCTCGAAGCTTCACTCATCATCTCGGAGAGCTCAAACTGG 721
DB 155 TTCACAGGAGCTCGAAGCTCGAAGCTTCACTCATCATCTCGGAGAGCTCAAACTGG 214
QY 722 CTGATTTTGGTCTTGCCCGGGCCAAAGTCATTTCCAGCCAGACATCTCTTCAGAAAGTCG 781
DB 215 CTGATTTTGGTCTTGCCCGGGCCAAAGTCATTTCCAGCCAGACATCTCTTCAGAAAGTCG 274
QY 782 TGACCTCTGGTATCCGGCCCCCTGATGCTTGTGCGAGCCACTGAATATTCCTTGAGC 841
DB 275 TGACCTCTGGTATCCGGCCCCCTGATGCTTGTGCGAGCCACTGAATATTCCTTGAGC 334
QY 842 TGGACATATGCGGGTGAAGGCTGATCTTATTAATGTTCCAGGGTCAACCTTTGTTTC 901
DB 335 TGGACATATGCGGGTGAAGGCTGATCTTATTAATGTTCCAGGGTCAACCTTTGTTTC 394
QY 902 CTGGGGTTTCCAAATCTTTGAAACAGCTGAGAAAATCTGGAGAGTCTTGAGAGTCCCTA 961
DB 395 CTGGGGTTTCCAAATCTTTGAAACAGCTGAGAAAATCTGGAGAGTCTTGAGAGTCCCTA 454
QY 962 CAGAGATATCTGGCGGGAGTCTCAAGCTACCTAATCAATCCAGAAATGTTCCAC 1021
DB 455 CAGAGATATCTGGCGGGAGTCTCAAGCTACCTAATCAATCCAGAAATGTTCCAC 514
QY 1022 TGGCTAAGCTGGAAGCTTCAATGTTGTCTGGAACAGGCTGGGAGAGTCTTGAAGCTG 1081
DB 515 TGGCTAAGCTGGAAGCTTCAATGTTGTCTGGAACAGGCTGGGAGAGTCTTGAAGCTG 574
QY 1082 AAGACCTGGGCTCCAGATGCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGAGAAAG 1141

DB 575 AAGACCTGGGCTCCAGATGCTAAAGGCTTTCCAGAGACCGGCTCTCCGCCAGAGAAAG 634
QY 1142 CACTTGTTCATGATTAATTTTCAAGCGCCCTGCAATCTCAGCTGTACCACTTCTGATGAG 1201
DB 635 CACTTGTTCATGATTAATTTTCAAGCGCCCTGCAATCTCAGCTGTACCACTTCTGATGAG 694
QY 1202 AGCTTTGTTTACAGTTTCAGAGTGAAGGCTAAAGCAGAAATGCTGACCTTTGGGCT 1261
DB 695 AGCTTTGTTTACAGTTTCAGAGTGAAGGCTAAAGCAGAAATGCTGACCTTTGGGCT 754
QY 1262 CCTACCAAGAAAGTCAACCAAGCCAGCTTAAAGCAATGCTGTGA 1308
DB 755 CCTACCAAGAAAGTCAACCAAGCCAGCTTAAAGCAATGCTGTGA 801

Search completed: December 26, 2004, 18:49:05
Job time : 382.27 secs

OM nucleic - nucleic search, using sw model

Run on: December 26, 2004, 17:39:26 ; Search time 3570.15 Seconds

17325.567 million cell updates/sec

Title: US-10-786-065-4

Sequence: 1 atcggtcacaagatctgtctgc.....agtctcagcaaatctgtctgtga 1308

Scoring table: IDENTITY_NUC

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

```
Minimum DB seq length: 0
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```

base :
GeneBml:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_ey:*
13: gb_un:*
14: gb_vl:*

```

Result No.	Query Match Score	Length	DB	ID	Description
1	100	100	1	1	...
2	95	95	2	2	...
3	90	90	3	3	...
4	85	85	4	4	...
5	80	80	5	5	...
6	75	75	6	6	...
7	70	70	7	7	...
8	65	65	8	8	...
9	60	60	9	9	...
10	55	55	10	10	...
11	50	50	11	11	...
12	45	45	12	12	...
13	40	40	13	13	...
14	35	35	14	14	...
15	30	30	15	15	...
16	25	25	16	16	...
17	20	20	17	17	...
18	15	15	18	18	...
19	10	10	19	19	...
20	5	5	20	20	...

1	1308	100.0	1308	6	AN265352	Sequence
2	1308	100.0	1308	6	AX571876	Sequence
3	1308	100.0	1308	6	AX772798	Sequence
4	1308	100.0	1584	6	AX772794	Sequence
5	1308	100.0	1628	6	AX772799	Sequence
6	1243.6	95.1	3675	9	HSM807533	Sequence
7	1200	91.7	2305	9	BC038807	Homo sapi
8	1185.4	90.6	1534	6	AX772791	Sequence
9	1185.4	90.6	1534	6	AX772795	Sequence
10	1185.4	90.6	1534	6	AB053308	Homo sapi
11	1009	77.1	2203	6	AR265350	Sequence
12	1009	77.1	2203	6	AX571873	Sequence
13	887.6	67.9	1376	6	AX644441	Sequence
14	873	66.7	1828	9	AK133512	Homo sapi
15	821	62.8	888	6	AX803407	Sequence
16	702.2	53.7	1077	6	AX056403	Sequence
17	481	36.8	481	6	AK175616	Sequence
18	481	36.8	481	6	AR236733	Sequence
19	439.6	33.6	4529	6	CQ77528	Sequence

20	439.6	33.6	44529	10	AF033655	AF033655	Mus musculus
21	438	33.5	26291	10	MM062311	U62391	Mus musculus
22	438	33.5	47132	10	AK192926	AK192926	Mus musculus
23	438	33.5	5127	10	BC068134	BC068134	Mus musculus
24	436.2	33.3	4957	6	C0776662	C0776662	Sequence
25	436.2	33.3	4957	6	AX928598	AX928598	Sequence
26	436.2	33.3	4957	6	AB020641	AB020641	Homo sapi
27	434.6	33.2	2043	9	AF119833	AF119833	Homo sapi
28	432.2	33.0	748	6	C0715966	C0715966	Sequence
29	404.2	30.9	2487	5	BC075148	BC075148	Xenopus la
30	401.6	30.7	1220	5	EX929373	EX929373	Gallus ga
31	399.2	30.5	1314	6	CQ71496	CQ71496	Sequence
32	398.4	28.9	1196	5	EX932969	EX932969	Gallus ga
33	339.4	25.9	2198	3	AF113367	AF113367	Drosophila
34	339.4	25.9	4789	3	BT010064	BT010064	Drosophila
35	337.8	25.8	2641	3	AF154403	AF154403	Drosophila
36	337.8	25.8	2755	3	AF154404	AF154404	Drosophila
37	337.8	25.8	2933	3	AF154402	AF154402	Drosophila
38	337.8	25.8	2942	3	AF154406	AF154406	Drosophila
39	337.8	25.8	2972	3	AF154401	AF154401	Drosophila
40	337.8	25.8	3065	3	AF154400	AF154400	Drosophila
41	337.8	25.8	3346	3	AF154405	AF154405	Drosophila
42	337.8	25.8	3617	3	AF153399	AF153399	Drosophila
43	337.8	25.8	3656	3	AF153398	AF153398	Drosophila
44	334.6	25.6	3526	3	DMPFRA19E	X93512	D.melanogaster
45	329.2	25.2	2703	6	CQ600711	CQ600711	Sequence

RESULT 1	AR265352	1308 bp	DNA	linear	PAT 10-APR-2003
LOCUS	AR265352				
DEFINITION	Sequence 4 from patent US 6492154.				
ACCESSION	AR265352				
VERSION	AR265352.1	GI:29693855			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1308)				
AUTHORS	van, C., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.				
TITLE	Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof				
JOURNAL	Patent: US 6492154-A 4 10-DEC-2002;				
FEATURES	Location/Qualifiers				
source	1..1308				

Query Match	100.0%;	Score 1308;	DB 6;	Length 1308;
Best Local Similarity	100.0%;	Pred NO. 0;		

	Desc	Local	Identity	Info	Mismatches	Indels	Gaps	
	Matches	1308	Conservative	0	0	0	0	0
QY	1	ATGGGTCAAGAGCTGTGTGCAAAAGACTGTACAGCTTGATGCAAGCTGTAACCAATTGTTTCA	60					
Db	1	ATGGGTCAAAAGCTGTGTGCAAAAGACTGTACAGCTTGATGCAAGCTGTAACCAATTGTTTCA	60					
QY	61	GAGGGAAGCGAGGACACAGCTGTGGAGAGAGTCAGGCTGAGACACGAGAGGCTGCGTTC	120					
Db	61	GAGGGAAGCGAGGACACAGCTGTGGAGAGAGTCAGGCTGAGACACGAGAGGCTGCGTTC	120					
QY	121	AAGCTAACAGACCTAAAGAAAGCATCATGTTCATCATTTCAACCCGAGGACCTT	180					
Db	121	AAGCTAACAGACCTAAAGAAAGCATCATGTTCATCATTTCAACCCGAGGACCTT	180					
QY	181	CAAGCTCCCGTGCCGAGAAAGTTCAAAAGTAAAGGACGAGATTAACAGTATTTGTTTT	240					
Db	181	CAAGCTCCCGTGCCGAGAAAGTTCAAAAGTAAAGGACGAGATTAACAGTATTTGTTTT	240					
QY	241	CAGGAGAAGATCTGAGGACAGGGTTTTTCAGTGGAGAAAGACCTTCCCTTTTGGGGCAGCC	300					

Db 241 CAGGAGAGATCTGAGGAGGGTTTTCACTGAGAGAGAGCTCCCTTTTGGGGAGCC 300
Qy 301 TCATCTTACTTGAACCTTGAGAGAGCTGGGTGAGAGCTTTATGCGACAGTTTCAAGAGG 360
Db 301 TCATCTTACTTGAACCTTGAGAGAGCTGGGTGAGAGCTTTATGCGACAGTTTCAAGAGG 360
Qy 361 ATTAGAGAAATAATGAGACACTAGTGGCTTTAAAAGTCACTGACATGAAATGCAAGGAA 420
Db 361 ATTAGAGAAATAATGAGACACTAGTGGCTTTAAAAGTCACTGACATGAAATGCAAGGAA 420
Qy 421 GGAGTCCCATTTACAGCTATCCGAGAGAGCTTCTCCTGAGAGGGTTTGAACATGCGAAT 480
Db 421 GGAGTCCCATTTACAGCTATCCGAGAGAGCTTCTCCTGAGAGGGTTTGAACATGCGAAT 480
Qy 481 ATTGTGCTCTCGATGACATAATCCACCAAGAGACCTGACATTCGTTTTGAATAC 540
Db 481 ATTGTGCTCTCGATGACATAATCCACCAAGAGACCTGACATTCGTTTTGAATAC 540
Qy 541 ATGCAACAAGACCTGGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCACTCTCAAT 600
Db 541 ATGCAACAAGACCTGGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCACTCTCAAT 600
Qy 601 GTCAAGCTTTTCAATGTTTCACTTTTGGGGGCTGGCTGACATCCACCAACACGTT 660
Db 601 GTCAAGCTTTTCAATGTTTCACTTTTGGGGGCTGGCTGACATCCACCAACACGTT 660
Qy 661 GTTCAAGAGACCTGAAACCTCAGAACTTATCATCAGTCACTGAGAGAGCTCAAACTG 720
Db 661 GTTCAAGAGACCTGAAACCTCAGAACTTATCATCAGTCACTGAGAGAGCTCAAACTG 720
Qy 721 GGTGATTTTGTCTTGGCCGGGGCAGTGCATTTCCAGCAGACATACCTTCAAGAAATC 780
Db 721 GGTGATTTTGTCTTGGCCGGGGCAGTGCATTTCCAGCAGACATACCTTCAAGAAATC 780
Qy 781 GTGACCTCTGTGATCCGGGCCCTGTATGCTTGTGAGAGCAGTGAATATTCCTGTAG 840
Db 781 GTGACCTCTGTGATCCGGGCCCTGTATGCTTGTGAGAGCAGTGAATATTCCTGTAG 840
Qy 841 CTGAGACATATGGGGTGCAGAGCTGCATCTTTATTTGAATGTTCCAGGGTCAACCTTTGTTT 900
Db 841 CTGAGACATATGGGGTGCAGAGCTGCATCTTTATTTGAATGTTCCAGGGTCAACCTTTGTTT 900
Qy 901 CCGGGGTTTCCAACATCCTTGAACAGCTGAGGAAATCTGGAGAGTGTGAGAGTCCCT 960
Db 901 CCGGGGTTTCCAACATCCTTGAACAGCTGAGGAAATCTGGAGAGTGTGAGAGTCCCT 960
Qy 961 ACAGAGATATCTGGCCGGGAGTCTTCAAGCTTACCTAACATCAGAAATGTTCCCA 1020
Db 961 ACAGAGATATCTGGCCGGGAGTCTTCAAGCTTACCTAACATCAGAAATGTTCCCA 1020
Qy 1021 CTGCTACGCCCTGAAAGCTTCACTGTTGTCTGAGAACAGCTGGGAGGGTCTCTGAAGCT 1080
Db 1021 CTGCTACGCCCTGAAAGCTTCACTGTTGTCTGAGAACAGCTGGGAGGGTCTCTGAAGCT 1080
Qy 1081 GAAAGACCTGGCCCTCCAGATGCTTAAAGGCTTTCCAGAGACCGAGTCTCCGCCAGGAA 1140
Db 1081 GAAAGACCTGGCCCTCCAGATGCTTAAAGGCTTTCCAGAGACCGAGTCTCCGCCAGGAA 1140
Qy 1141 GCACTTGTTCATGATTAATTTTCAAGCGCTGCGCATCTCAGCTGTACCAAGCTTCTGATGAG 1200
Db 1141 GCACTTGTTCATGATTAATTTTCAAGCGCTGCGCATCTCAGCTGTACCAAGCTTCTGATGAG 1200
Qy 1201 GAGTCTTGTTCATGATTTCAAGAGTGAAGCTTAAAGCCGAAATGTGTGACCTTTTGGCC 1260
Db 1201 GAGTCTTGTTCATGATTTCAAGAGTGAAGCTTAAAGCCGAAATGTGTGACCTTTTGGCC 1260
Qy 1261 TCTTACCAAGAGTCAACCAAGCCAGTTTGAAGAAATGCTGTGGA 1308
Db 1261 TCTTACCAAGAGTCAACCAAGCCAGTTTGAAGAAATGCTGTGGA 1308

RESULT 2

AX571876
LOCUS AX571876 1308 bp DNA linear PAT 29-NOV-2002
DEFINITION Sequence 4 from Patent WO02061060.
ACCESSION AX571876
VERSION AX571876.1 GI:26004001
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 1308
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 1308; DB 6; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGTCAAAGAGCTGTGTGCAAGAGCTGTACAGCTGTGATGAGCTGTACCATTTGTCA 60
Db 1 ATGGGTCAAAGAGCTGTGTGCAAGAGCTGTACAGCTGTGATGAGCTGTACCATTTGTCA 60
Qy 61 GAGGAGGCGAGGACACACAGCTGTGTGAGAGTCAAGCTGTGAGACCAAGAGGCTGCTTC 120
Db 61 GAGGAGGCGAGGACACACAGCTGTGTGAGAGTCAAGCTGTGAGACCAAGAGGCTGCTTC 120
Qy 121 AAGCTAACAGACCTTAAAGAGCATCATGTTCCATGACTTCATTTCACCCAGGGACTT 180
Db 121 AAGCTAACAGACCTTAAAGAGCATCATGTTCCATGACTTCATTTCACCCAGGGACTT 180
Qy 181 CAAGTCCCGCTGCTCCAGAGTTCAAGAGTAAAGGCCACGAGTAAAGATGTTGTTT 240
Db 181 CAAGTCCCGCTGCTCCAGAGTTCAAGAGTAAAGGCCACGAGTAAAGATGTTGTTT 240
Qy 241 CAGGAAGAGATCTAGAGCAGAGGTTTTTCAAGAGAGAGAGCTCCCTTTTGGGGCAGCC 300
Db 241 CAGGAAGAGATCTAGAGCAGAGGTTTTTCAAGAGAGAGAGCTCCCTTTTGGGGCAGCC 300
Qy 301 TCATCTTACTTGAACCTTGAGAGAGCTGGGTGAGAGCTTTATGCGACAGTTTCAAGAGG 360
Db 301 TCATCTTACTTGAACCTTGAGAGAGCTGGGTGAGAGCTTTATGCGACAGTTTCAAGAGG 360
Qy 361 ATTAGCAAGATTAATGAGCACTAGTGGCTTTAAAAGTCACTGACATGAAATGCAAGGAA 420
Db 361 ATTAGCAAGATTAATGAGCACTAGTGGCTTTAAAAGTCACTGACATGAAATGCAAGGAA 420
Qy 421 GGAGTCCCATTTACAGCTATCCGAGAGAGCTTCTCCTGAGAGGGTTTGAACATGCGAAT 480
Db 421 GGAGTCCCATTTACAGCTATCCGAGAGAGCTTCTCCTGAGAGGGTTTGAACATGCGAAT 480
Qy 481 ATTGTGCTCTCGATGACATAATCCACCAAGAGACCTGACATTCGTTTTGAATAC 540
Db 481 ATTGTGCTCTCGATGACATAATCCACCAAGAGACCTGACATTCGTTTTGAATAC 540
Qy 541 ATGCAACAAGACCTGGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCACTCTCAAT 600
Db 541 ATGCAACAAGACCTGGGCCAGTATATGTCTCAGCATCCAGAGGGCTTCACTCTCAAT 600
Qy 601 GTCAAGCTTTTCAATGTTTCACTTTTGGGGGCTGGCTGACATCCACCAACACGTT 660
Db 601 GTCAAGCTTTTCAATGTTTCACTTTTGGGGGCTGGCTGACATCCACCAACACGTT 660
Qy 661 GTTCAAGAGACCTGAAACCTCAGAACTTATCATCAGTCACTGAGAGAGCTCAAACTG 720
Db 661 GTTCAAGAGACCTGAAACCTCAGAACTTATCATCAGTCACTGAGAGAGCTCAAACTG 720

Db 661 CTTACAGGAGACCTGAAACCTTCACTATCATGACCTGGAGAGCTCAACTG 720
Qy 721 GCTGATTTTGTCTTCCCGGGCCAACTGCTATCCAGCCAGACATATCTTCAGAACTC 780
Db 721 GCTGATTTTGTCTTCCCGGGCCAACTGCTATCCAGCCAGACATATCTTCAGAACTC 780
Qy 781 GTGACCTCTGTGATCCGGCCCCCTGATGCTTTTCTGGAGGCCAGTGAATTTCTCTGAG 840
Db 781 GTGACCTCTGTGATCCGGCCCCCTGATGCTTTTCTGGAGGCCAGTGAATTTCTCTGAG 840
Qy 841 CTGGAATATGAGGGGTGAGGCTGATCTTTATTTGAATGTTCCAGGGTCAACCTTTGTTT 900
Db 841 CTGGAATATGAGGGGTGAGGCTGATCTTTATTTGAATGTTCCAGGGTCAACCTTTGTTT 900
Qy 901 CCGGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
Db 901 CCGGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
Qy 961 ACAGAGGATATCTTGGCCGGAGTCTCCAAAGCTACCTAACATCCAGATGCTTCCCA 1020
Db 961 ACAGAGGATATCTTGGCCGGAGTCTCCAAAGCTACCTAACATCCAGATGCTTCCCA 1020
Qy 1021 CTGCTTACGCTCGAAGCTTCACTGTTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
Db 1021 CTGCTTACGCTCGAAGCTTCACTGTTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
Qy 1081 GAAGACCTGGCTCCAGATGCTTAAAGGCTTCCAGAGACCGGCTCTCCGCCAGAGAA 1140
Db 1081 GAAGACCTGGCTCCAGATGCTTAAAGGCTTCCAGAGACCGGCTCTCCGCCAGAGAA 1140
Qy 1141 GCACTTTGTTCAATATTTTCAAGCCCTGCTCCATCTCAAGTGTACCAAGTCTTCTGATGAG 1200
Db 1141 GCACTTTGTTCAATATTTTCAAGCCCTGCTCCATCTCAAGTGTACCAAGTCTTCTGATGAG 1200
Qy 1201 GAGTCTTTGTTCAATATTTTCAAGAGTGAAGCTTAAAGCAAGAAATGTTGACCTTTTGACC 1260
Db 1201 GAGTCTTTGTTCAATATTTTCAAGAGTGAAGCTTAAAGCAAGAAATGTTGACCTTTTGACC 1260
Qy 1261 TCTTACAGAAAGGTGACCAAGCCCAAGTTTGAAGAAATGCTGTGTA 1308
Db 1261 TCTTACAGAAAGGTGACCAAGCCCAAGTTTGAAGAAATGCTGTGTA 1308

RESULT 3
AX772798 1308 bp DNA linear PAT 09-JUL-2003
LOCUS Sequence 8 from Patent WO03046167.
DEFINITION AX772798
ACCESSION AX772798
VERSION AX772798.1 GI:32485025
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
Bayer Aktiengesellschaft (DE)
Location/Qualifiers
1. .1308
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="serine/threonine protein kinase"

FEATURES
source

ORIGIN

Query Match 100.0%; Score 1308; DB 6; Length 1308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGTCAGAGCTGTGTGCAAGACTGTACAGCTGTGATGAGCTGCTTACCTTGTCA 60
|||||
Db 1 ATGGGTCAGAGCTGTGTGCAAGACTGTACAGCTGTGATGAGCTGCTTACCTTGTCA 60

Db 1 ATGGGTCAGAGCTGTGTGCAAGACTGTACAGCTGTGATGAGCTGCTTACCTTGTCA 60
Qy 61 GAGGAGGCGAGGACACAGCTGTGAGAGTCAAGCTTGAAGCAACGAGGCTGCTTC 120
Db 61 GAGGAGGCGAGGACACAGCTGTGAGAGTCAAGCTTGAAGCAACGAGGCTGCTTC 120
Qy 121 AAGCTTACAGACTTAAAGAGATCATGTTCTCAATGACTTCAATTTCAACCCAGGGAGCTT 180
Db 121 AAGCTTACAGACTTAAAGAGATCATGTTCTCAATGACTTCAATTTCAACCCAGGGAGCTT 180
Qy 181 CAAGCTGCCGTCAGAGAGTTCAAGAGTAAAGGCCAGGAGTAAAGTATTTT 240
Db 181 CAAGCTGCCGTCAGAGAGTTCAAGAGTAAAGGCCAGGAGTAAAGTATTTT 240
Qy 241 CAGAAAGAGATCTGAGGAGGGTTTCACTGAGAGAGAGGCTCTCTTTGGGAGAGCC 300
Db 241 CAGAAAGAGATCTGAGGAGGGTTTCACTGAGAGAGAGGCTCTCTTTGGGAGAGCC 300
Qy 301 TCATCTTACTTGAATCTTGAAGAGCTGGTGAAGGCTTTATGCAACATTTTCAAGGGG 360
Db 301 TCATCTTACTTGAATCTTGAAGAGCTGGTGAAGGCTTTATGCAACATTTTCAAGGGG 360
Qy 361 ATTAGAGAAATTAATGGAACAATGAGCTTTAAAGTATGATGATGATGATGATGATGAT 420
Db 361 ATTAGAGAAATTAATGGAACAATGAGCTTTAAAGTATGATGATGATGATGATGATGAT 420
Qy 421 GAGATCCATTTTCAAGCTATCCGAGAGGCTTCTCTGAGAGGTTTGAACATGCAAT 480
Db 421 GAGATCCATTTTCAAGCTATCCGAGAGGCTTCTCTGAGAGGTTTGAACATGCAAT 480
Qy 481 ATTGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 481 ATTGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 541 ATGACACAGAGCTGAGCCAGATATATGCTCAAGATCCAGAGGAGGCTTATCTCTAAT 600
Db 541 ATGACACAGAGCTGAGCCAGATATATGCTCAAGATCCAGAGGAGGCTTATCTCTAAT 600
Qy 601 GTGAGCTTTTCAATGTTTCACTTTTGGGGCTGAGCTGATCCAGCAACCAACGTT 660
Db 601 GTGAGCTTTTCAATGTTTCACTTTTGGGGCTGAGCTGATCCAGCAACCAACGTT 660
Qy 661 CTTACAGGAGACTGAAACCTTCAAGATCTTCAATGATGATGATGATGATGATGATGATGAT 720
Db 661 CTTACAGGAGACTGAAACCTTCAAGATCTTCAATGATGATGATGATGATGATGATGATGAT 720
Qy 721 GCTGATTTTGTCTTGGCCGGGCAAGTCAATTTCCAGCAGACATCTTCCAGAACTC 780
Db 721 GCTGATTTTGTCTTGGCCGGGCAAGTCAATTTCCAGCAGACATCTTCCAGAACTC 780
Qy 781 GTGACCTCTGTGATCCGAGCCCTGATGCTTTGCTGGAGCACTGAATATTTCTCTGAG 840
Db 781 GTGACCTCTGTGATCCGAGCCCTGATGCTTTGCTGGAGCACTGAATATTTCTCTGAG 840
Qy 841 CTGGAATATGAGGGGTGAGGCTGATCTTTATTTGAATGTTCCAGGGTCAACCTTTGTTT 900
Db 841 CTGGAATATGAGGGGTGAGGCTGATCTTTATTTGAATGTTCCAGGGTCAACCTTTGTTT 900
Qy 901 CCGGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
Db 901 CCGGGGTTTCCAACTCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCT 960
Qy 961 ACAGAGGATATCTTGGCCGGAGTCTCCAAAGCTACCTAACATCCAGATGCTTCCCA 1020
Db 961 ACAGAGGATATCTTGGCCGGAGTCTCCAAAGCTACCTAACATCCAGATGCTTCCCA 1020
Qy 1021 CTGCTTACGCTCGAAGCTTCACTGTTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
Db 1021 CTGCTTACGCTCGAAGCTTCACTGTTCTGGAACAGGCTGGGCAAGGTTCTGAAAGCT 1080
Qy 1081 GAAGACCTGGCTCCAGATGCTTAAAGGCTTCCAGAGACCGGCTCTCCGCCAGAGAA 1140
Db 1081 GAAGACCTGGCTCCAGATGCTTAAAGGCTTCCAGAGACCGGCTCTCCGCCAGAGAA 1140

QY 1141 GCACCTGTTATGATTTATTTTCAGGCGCCCTGCGCATCTTCAGCTGTACCAAGCTTCTGATGAG 1200
DB 1141 GCACCTGTTATGATTTATTTTCAGGCGCCCTGCGCATCTTCAGCTGTACCAAGCTTCTGATGAG 1200
QY 1201 GAGCTCTTTGTTTACAGTTTCAGAGAGTGAAGGCTAAAGCCAGAAATGTGTGACTTTTGGCC 1260
DB 1201 GAGCTCTTTGTTTACAGTTTCAGAGAGTGAAGGCTAAAGCCAGAAATGTGTGACTTTTGGCC 1260
QY 1261 TCCTACCAAGAGGTCCACCAAGCCCAAGTTTACCAATGTCTGTGA 1308
DB 1261 TCCTACCAAGAGGTCCACCAAGCCCAAGTTTACCAATGTCTGTGA 1308
RESULT 4
AX772794 1584 bp DNA linear PAT 09-JUL-2003
LOCUS Sequence 4 from Patent WO03046167.
DEFINITION AX772794
ACCESSION AX772794
VERSION AX772794.1 GI:32485023
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Koehler, R.H.
TITLE Regulation of human serine/threonine protein kinase
JOURNAL Patent: WO 03046167-A 4 05-JUN-2003;
Bayer Aktiengesellschaft (DE)
LOCATION/Qualifiers
FEATURES
source 1. 1584
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="serine/threonine protein kinase"
ORIGIN
Query Match 100.0%; Score 1308; DB 6; Length 1584;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ANGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTGTGATGAGCTGCTACCATTTGTTCA 60
DB 1 ANGGGTCAAGAGCTGTGTGCAAGACTGTACAGCTGTGATGAGCTGCTACCATTTGTTCA 60
QY 61 GAGGAGGCGAGGACACACAGCTGTGCGAGAGTCAAGCTGTGAGACCAAGAGGCTGCGTTC 120
DB 61 GAGGAGGCGAGGACACACAGCTGTGCGAGAGTCAAGCTGTGAGACCAAGAGGCTGCGTTC 120
QY 121 AAGCTTAACAGACCTTAAGAAAGATCATGTTCCATGACTTCACTTCAACCCAGAGGACTT 180
DB 121 AAGCTTAACAGACCTTAAGAAAGATCATGTTCCATGACTTCACTTCAACCCAGAGGACTT 180
QY 181 CAAGTCCCGTGTCCAGAAAGTCAAGAGTAAAGGCCAGAGTAACAGTATGTTTT 240
DB 181 CAAGTCCCGTGTCCAGAAAGTCAAGAGTAAAGGCCAGAGTAACAGTATGTTTT 240
QY 241 CAGGAAGAGATCTGAGGAGGAGGTTTCAAGTGAAGAGAGCTCCCTTTTGGGAGGCC 300
DB 241 CAGGAAGAGATCTGAGGAGGAGGTTTCAAGTGAAGAGAGCTCCCTTTTGGGAGGCC 300
QY 301 TCATCTTACTTGAACCTTGGAAGGCTGAGGCTGTTAATGCCAGATTTTCAAGAGGAG 360
DB 301 TCATCTTACTTGAACCTTGGAAGGCTGAGGCTGTTAATGCCAGATTTTCAAGAGGAG 360
QY 361 ATTAGAGAAATAATGAGCAACTAGTGGCTTTAAAGTATCGAGTGAATGAGAGGAA 420
DB 361 ATTAGAGAAATAATGAGCAACTAGTGGCTTTAAAGTATCGAGTGAATGAGAGGAA 420
QY 421 GGAGTCCCATTTTACAGCTATCCAGAGAGCTTCTCTCTGAAGGTTTGAACAATGCCAAT 480
DB 421 GGAGTCCCATTTTACAGCTATCCAGAGAGCTTCTCTCTGAAGGTTTGAACAATGCCAAT 480

QY 481 ATTGTGCTCTGCTGATGACATAATCCACACCAAGAGACCTGACATTCGTTTTGAAATAC 540
DB 481 ATTGTGCTCTGCTGATGACATAATCCACACCAAGAGACCTGACATTCGTTTTGAAATAC 540
QY 541 ATGCACACAGACCTGGCCAGATATATGTCTCAGCATCCAGAGAGGGCTTATCTCTAAT 600
DB 541 ATGCACACAGACCTGGCCAGATATATGTCTCAGCATCCAGAGAGGGCTTATCTCTAAT 600
QY 601 GTGAGCTTTTCAATGTTTCACTTTTGCAGGAGCTGTGATCATCCACCAACACAGTT 660
DB 601 GTGAGCTTTTCAATGTTTCACTTTTGCAGGAGCTGTGATCATCCACCAACACAGTT 660
QY 661 CTTTACAGAGGACCTGAAACCTCAGAACTTACTATCATGCTACCTGGAGAGCTCAACTG 720
DB 661 CTTTACAGAGGACCTGAAACCTCAGAACTTACTATCATGCTACCTGGAGAGCTCAACTG 720
QY 721 GCTGATTTTGTCTTGGCCCGGAGCAAGTCCATTTCCAGCAGACATPACTTTCAGAGTTC 780
DB 721 GCTGATTTTGTCTTGGCCCGGAGCAAGTCCATTTCCAGCAGACATPACTTTCAGAGTTC 780
QY 781 GTGACCTCTGTGTACCGGCCCCCTGATGTCTTGTGAGAGCCATGAAATATCTCTGAG 840
DB 781 GTGACCTCTGTGTACCGGCCCCCTGATGTCTTGTGAGAGCCATGAAATATCTCTGAG 840
QY 841 CTGACATATGTGGGTGTCAGAGCTGCATCTTTATTTGAATGTTCAGAGGTCAACTTTGTT 900
DB 841 CTGACATATGTGGGTGTCAGAGCTGCATCTTTATTTGAATGTTCAGAGGTCAACTTTGTT 900
QY 901 CCGGGGTTTCCAAATCCTTTGAAACAGCTGAGAAATCTGGAGAGTGTGAGAGTCCCT 960
DB 901 CCGGGGTTTCCAAATCCTTTGAAACAGCTGAGAAATCTGGAGAGTGTGAGAGTCCCT 960
QY 961 ACAGAGATATCTTGGCCGAGAGTCTCAAGCTTAACTCAATTCAGAAATGTTTCCA 1020
DB 961 ACAGAGATATCTTGGCCGAGAGTCTCAAGCTTAACTCAATTCAGAAATGTTTCCA 1020
QY 1021 CTGCTACGCTCTGAAAGCTTATGTTCTGTGAAACAGCTGGGAGAGTTCTTAAGCT 1080
DB 1021 CTGCTACGCTCTGAAAGCTTATGTTCTGTGAAACAGCTGGGAGAGTTCTTAAGCT 1080
QY 1081 GAAAGACTGGCTCTCCAGATGCTAAAGGCTTCCCAAGACCGGCTCTCCGCCAGAA 1140
DB 1081 GAAAGACTGGCTCTCCAGATGCTAAAGGCTTCCCAAGACCGGCTCTCCGCCAGAA 1140
QY 1141 GCATTTGTTCAATGATTTATTCAGGCGCCCTGCGCATCTCAAGCTTCTGATGAG 1200
DB 1141 GCATTTGTTCAATGATTTATTCAGGCGCCCTGCGCATCTCAAGCTTCTGATGAG 1200
QY 1201 GAGTCTTTGTTTACAGTTTCAAGAGTGAAGGCTTAAAGCCAGAAATGTGTGACTTTTGGCC 1260
DB 1201 GAGTCTTTGTTTACAGTTTCAAGAGTGAAGGCTTAAAGCCAGAAATGTGTGACTTTTGGCC 1260
QY 1261 TCCTACCAAGAGGTCCACCAAGCCCAAGTTTACCAATGTGTGTGA 1308
DB 1261 TCCTACCAAGAGGTCCACCAAGCCCAAGTTTACCAATGTGTGTGA 1308
RESULT 5
AX772799 1628 bp DNA linear PAT 09-JUL-2003
LOCUS Sequence 9 from Patent WO03046167.
DEFINITION AX772799
ACCESSION AX772799
VERSION AX772799.1 GI:32485026
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Koehler, R.H.
TITLE Regulation of human serine/threonine protein kinase
JOURNAL Patent: WO 03046167-A 9 05-JUN-2003;

ORIGIN

MALCO

99

Db

SOURCE

COMMENT

00000000

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"
```

/clone="DKFP686K20144"
/tissue_type="human colon endothel primary cell culture"
/clone_id="686 (synonym: h1cc3). Vector pSPORT1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
3589.3594
polYA_signal
polYA_site
3612
ORIGIN

Query Match 95.1%; Score 1243.6; DB 9; Length 3675;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 1285; Conservative 0; Mismatches 4; Indels 19; Gaps 2;

QY 1 ATGGGCTCAAGAGCTGTGTGCAAGAGCTGTACAGCTGTGATGACAGCTGTCTACATTTGTTCA 60
DB 51 ATGGGCTCAAGAGCTGTGTGCAAGAGCTGTACAGCTGTGATGACAGCTGTCTACATTTGTTCA 110
QY 61 GAGGAGGCGAGGACACACAGCTGTGCGAGGAGTCAAGCTGTGACCAAGAGGCTGCGTTG 120
DB 111 GAGGAGGCGAGGACACACAGCTGTGCGAGGAGTCAAGCTGTGACCAAGAGGCTGCGTTG 170
QY 121 AAGCTTACAGACCTTAAAGAGATCATGTTTCATTTCACTTCACTTCACTTCACTTCACTT 180
DB 171 AAGCTTACAGACCTTAAAGAGATCATGTTTCATTTCACTTCACTTCACTTCACTTCACTT 230
QY 181 CAAGCTGCGCGTCCAGAGATTCAGAGTAAAGGCGACAGGATTAAGTATTTGTTT 240
DB 231 CAAGCTGCGCGTCCAGAGATTCAGAGTAAAGGCGACAGGATTAAGTATTTGTTT 290
QY 241 CAGAGAGAGATCTGAGGCGAGGTTTTCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 291 CAGAGAGAGATCTGAGGCGAGGTTTTCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 350
QY 301 TCATCTTACTTGAACCTTGAAGAGCTGCGTGAAGGCTTTATGCGACATTTCAAGAGG 360
DB 351 TCATCTTACTTGAACCTTGAAGAGCTGCGTGAAGGCTTTATGCGACATTTCAAGAGG 410
QY 361 ATTAGAGAGATTAATGAGACACTAGTGGCTTTAAAGTCACTGACATTAATGAGAGAG 420
DB 411 ATTAGAGAGATTAATGAGACACTAGTGGCTTTAAAGTCACTGACATTAATGAGAGAG 470
QY 421 GAGGCTCCATTTACAGCTATCCGAGAGAGCTTCTCTCTGAGAGGTTTGAAGAGAGAG 480
DB 471 GAGGCTCCATTTACAGCTATCCGAGAGAGCTTCTCTCTGAGAGGTTTGAAGAGAGAG 530
QY 481 ATTTGCTCTCTGATGACATTAATCCACCAAGAGAGACTGACATTTGTTTGAATAC 540
DB 531 ATTTGCTCTCTGATGACATTAATCCACCAAGAGAGACTGACATTTGTTTGAATAC 590
QY 541 ATGACACAGAGAGCTGGGCGGAGTATGTCACAGATCCAGAGAGGCTTCACTCTCAAT 600
DB 591 ATGACACAGAGAGCTGGGCGGAGTATGTCACAGATCCAGAGAGGCTTCACTCTCAAT 650
QY 601 GTGAGACTTTTCACTGTTTCACTTTTTCGCGGCGCTGCGTATCCACCAAGAGAGAG 660
DB 651 GTGAGACTTTTCACTGTTTCACTTTTTCGCGGCGCTGCGTATCCACCAAGAGAGAG 710
QY 661 CTTTCAAGAGAGAGCTGAAAGCTGAACTTACTATGATGATCACTGCGAGAGCTCAAC 720
DB 711 CTTTCAAGAGAGAGCTGAAAGCTGAACTTACTATGATGATCACTGCGAGAGCTCAAC 770
QY 721 GGTGATTTTGTGTTGCGCGGCGGAGTCACTTCCAGAGAGATTAATCTTCAAGAG 780
DB 771 GGTGATTTTGTGTTGCGCGGCGGAGTCACTTCCAGAGAGATTAATCTTCAAGAG 830
QY 781 GTGAGACCTCTGTATCGCGGCGGCTGTATGCTTTGCTGAGAGCACTGAATTTCTCTGAG 840
DB 831 GTGAGACCTCTGTATCGCGGCGGCTGTATGCTTTGCTGAGAGCACTGAATTTCTCTGAG 890
QY 841 CTGGAACATATGGGCTGAGGCTGATCTTTTATTTGAATGTTTCCAGAGGCTCAACTTTGTTT 900
DB 891 CTGGAACATATGGGCTGAGGCTGATCTTTTATTTGAATGTTTCCAGAGGCTCAACTTTGTTT 950

QY 901 CTTGGGCTTTCACACATCTCTTGAACAGCTGAGAGAAATCTGGAGAGTGTCTGGAGTCCCT 960
DB 951 CCT-GGGTTTCCAAACATCTTTGAACAGCTGAGAGAAATCTGGAGAGTGTCTGGAGTCCCT 1009
QY 961 ACAGAGATATCTTGGCGGAGAGCTTCCAGTACCTTAACCTAATTCAGAAATGTTTCCCA 1020
DB 1010 ACAGAGATATCTTGGCGGAGAGCTTCCAGTACCTTAACCTAATTCAGAAATGTTTCCCA 1069
QY 1021 CTGCTTACGCTTGAAGCTTCACTGTTTCTGTGAACAGGCTGGGAGGTTCTTGAAGCT 1080
DB 1070 CTGCTTACGCTTGAAGCTTCACTGTTTCTGTGAACAGGCTGGGAGGTTCTTGAAGCT 1129
QY 1081 GAAGACCTGCGCTCCAGATGCTAAAGGCTTCCAGAGACCGGCTCCGCGCAGGAA 1140
DB 1130 GAAGACCTGCGCTCCAGATGCTAAAGGCTTCCAGAGACCGGCTCCGCGCAGGAA 1189
QY 1141 GCACTTGTTCATGATTAATTTACAGCGGCTTCCATCTCAGCTGTACAGCTTCTGTAG 1200
DB 1190 GCACTTGTTCATGATTAATTTACAGCGGCTTCCATCTCAGCTGTACAGCTTCTGTAG 1246
QY 1201 GAGTCTTGTTTACAGTTCAGAGATGAGGCTTAAGCCGAAATGTTGATCCTTTGGCC 1260
DB 1247 -----GTTTCAGAGAGTGAAGCTTAAGCCGAAATGTTGATCCTTTGGCC 1291
QY 1261 TCCTACAGAGAGATCACACCCAGCCAGTTTAGCAATGCTGGTGA 1308
DB 1292 TCCTACAGAGAGATCACACCCAGCCAGTTTAGCAATGCTGGTGA 1339

RESULT 7
LOCUS BC038807
DEFINITION Homo sapiens amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 7, mRNA (cdna clone MGC:46456 IMAGE:5200755), complete cds.
ACCESSION BC038807
VERSION BC038807.1 GI:24416556
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2305)
Strusberg R.L., Pelngold E.A., Gronow L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhac N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquelliano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Young A.C., Shevchenko Y., Sanchez A., Whitting M., Madan A., Touchman J.W., Green E.D., Bouffard G.G., Blakeley R.W., Grimwood J., Schmutz J., Myers R.M., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smallwood D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 2305)
JOURNAL MEDLINE
PUBMED 12477932
REFERENCE 2388257
AUTHORS Strausberg R.
TITLE Direct Submission
SUBMITTED (25-OCT-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhari, N., Aylee, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Guan, X., Gupta, J., Haghighi, P., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaardi, R., Madero, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stancitop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAK Plate: 78 Row: f Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21040234.

FEATURES
source

Location/Qualifiers
1. 2305
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:46456 IMAGE:5200755"
/cissue_type="Brain, adult, 6 pooled whole brains"
/clone_11b="NIH_MGC_114"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 2305
/gene="ALS2CR7"
/db_xref="LocusID:65061"
198. 1247
/gene="ALS2CR7"
/product="ALS2CR7 protein"
/protein_id="AAH38807.1"
/db_xref="GI:24416557"
/db_xref="LocusID:65061"
/translation="MTSFRRLQAAARQFKSRKPRNSDCFOEEDLRQGFOMKSL
PGAASVYLNLEKLGESYATVYKGISRINQOLVALKVISMAABEVPPTAIRBASIL
KGLKHANVILHDIHTKETTLPFEYEMHTDIAOYMSCHPGLHPHNPVLPFOILRG
LAYIHGHVLRHDLKRONLILSHLGEKLADRLAKRSIPQOTISSEVYVITMYRPD
ALGATVSSSELDINGAGCTFLEMFQGFVPGVSNITELQEKIWEVLGVPTEDTPG
VSKLPNYPNEMFPLPTPRSLHVVMNRLGVPBAEDLASQMLKGFPRDRVASQALVHD
YFSALPSQLYQLPDE"
351. 1205
/gene="ALS2CR7"
/note="S_TKC; Region: Serine/Threonine protein kinases,
catalytic domain"
/db_xref="CD:smart00220"

ORIGIN

Query Match 91.7%; Score 1200; DB 9; Length 2305;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCAAGAGCTGTGTCAGAACTGTACAGCCCTGATGAGCTGTACATTTGTTCA 60
DB 45 ATGGGTCAAGAGCTGTGTCAGAACTGTACAGCCCTGATGAGCTGTACATTTGTTCA 104
QY 61 GAGGAGGCGGAGGACACAGCTGTGCGAGAGTCAAGCTGAGACCAACGAGGCTGCGTTTC 120
DB 105 GAGGAGGCGGAGGACACAGCTGTGCGAGAGTCAAGCTGAGACCAACGAGGCTGCGTTTC 164
QY 121 AAGCTAAGACGACCTAAGAAAGATCATGTTCCATTCATTACACCCGAGGAGCTT 180
|||||

DB 165 AAGCTAAGACGACCTAAGAAAGATCATGTTCCATTCATTACACCCGAGGAGCTT 224
QY 181 CAAGTCCCGCTGCCGAGAGTTCAAGATAAAGCCACGAGTAACAGTATTTT 240
DB 225 CAAGTCCCGCTGCCGAGAGTTCAAGATAAAGCCACGAGTAACAGTATTTT 284
QY 241 CAGGAAGAGATCTGAGGAGGTTTTCAGTGTGAGAGAAAGCTCCCTTTTGGGAGGC 300
DB 285 CAGGAAGAGATCTGAGGAGGTTTTCAGTGTGAGAGAAAGCTCCCTTTTGGGAGGC 344
QY 301 TCATCTTACTTGAACCTTGAGAGAGCTGAGGAGCTCTTATGCGACAGTTTCAAGGAG 360
DB 345 TCATCTTACTTGAACCTTGAGAGAGCTGAGGAGCTCTTATGCGACAGTTTCAAGGAG 404
QY 361 ATTAGCAATATATATGACCACTAGTGGCTTTTAAAGTCATGACATGATGACAGAGAA 420
DB 405 ATTAGCAATATATATGACCACTAGTGGCTTTTAAAGTCATGACATGATGACAGAGAA 464
QY 421 GGAATCCCATTTACAGCTATCCGAGAGGCTTCTCTGAGAGGTTTGAATGCGCAT 480
DB 465 GGAATCCCATTTACAGCTATCCGAGAGGCTTCTCTGAGAGGTTTGAATGCGCAT 524
QY 481 ATTGAGCTCTGATGACATTAATCCACCAAGAGACCTGACATTTGTTTGAATAC 540
DB 525 ATTGAGCTCTGATGACATTAATCCACCAAGAGACCTGACATTTGTTTGAATAC 584
QY 541 ATGACACAGACCTGAGCCAGATATATGTCAGCATCCAGAGGCTTCATCCTCAAT 600
DB 585 ATGACACAGACCTGAGCCAGATATATGTCAGCATCCAGAGGCTTCATCCTCAAT 644
QY 601 GTGACATTTTATATGTTTCACTTTTGGGGGCTGAGCTATCATCCACCAACAGCTT 660
DB 645 GTGACATTTTATATGTTTCACTTTTGGGGGCTGAGCTATCATCCACCAACAGCTT 704
QY 661 CTTCACAGGAGCTGAGAACTCAGAACTCATCATGACCTGAGAGAGCTCAAACTG 720
DB 705 CTTCACAGGAGCTGAGAACTCAGAACTCATCATGACCTGAGAGAGCTCAAACTG 764
QY 721 GCTGATTTTGTGCTTGGCCCGGAGCAAGTTCATTCACAGCAATCTTTCAGAGTC 780
DB 765 GCTGATTTTGTGCTTGGCCCGGAGCAAGTTCATTCACAGCAATCTTTCAGAGTC 824
QY 781 GTGACCTCTGTGATCCGCGCCCTGATGCTTGTGCGGAGCACTGATATTTCTTGAG 840
DB 825 GTGACCTCTGTGATCCGCGCCCTGATGCTTGTGCGGAGCACTGATATTTCTTGAG 884
QY 841 CTGACATATGAGGAGTGCAGGCTGCATCTTATTTGAATGTTCCAGGTCACCTTGT 900
DB 885 CTGACATATGAGGAGTGCAGGCTGCATCTTATTTGAATGTTCCAGGTCACCTTGT 944
QY 901 CTTGAGGCTTCCAGATCTTGAACAGCTGAGAAATCTGAGAGGTGCTGAGAGTCCCT 960
DB 945 CTTGAGGCTTCCAGATCTTGAACAGCTGAGAAATCTGAGAGGTGCTGAGAGTCCCT 1004
QY 961 AAGAGATATCTTGGCCCGGAGATCTTCCAGATCTTCAATCAATTCAGATGTTCCA 1020
DB 1005 AAGAGATATCTTGGCCCGGAGATCTTCCAGATCTTCAATCAATTCAGATGTTCCA 1064
QY 1021 CTGCTAGAGCTGAGAGCTTATGATGTTGTTGAGAAAGGCTGAGGAGGTTCTGAGAGCT 1080
DB 1065 CTGCTAGAGCTGAGAGCTTATGATGTTGTTGAGAAAGGCTGAGGAGGTTCTGAGAGCT 1124
QY 1081 GAAAGCTGAGCTTCCAGATGCTTAAAGGCTTTCCAGAGACCGGCTCTCCGCGCAGAA 1140
DB 1125 GAAAGCTGAGCTTCCAGATGCTTAAAGGCTTTCCAGAGACCGGCTCTCCGCGCAGAA 1184
QY 1141 GCACTTTGATGATATTTTACGCGCTGCGCATCTCAGCTGTACAGCTTCTGATGAG 1200
DB 1185 GCACTTTGATGATATTTTACGCGCTGCGCATCTCAGCTGTACAGCTTCTGATGAG 1244
RESULT 8
AX772791

LOCUS	AX772791	1534 bp	mRNA	linear	PAT 09-JUL-2003
DEFINITION	Sequence 1 from Patent WO03046167.				
ACCESSION	AX772791				
VERSION	AX772791.1	GI:32485021			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Kochler, R. H.				
TITLE	Regulation of human serine/threonine protein kinase				
JOURNAL	Patent: WO 03046167-A 1 05-JUN-2003;				
	Bayer Aktiengesellschaft (DE)				
FEATURES	Location/Qualifiers				
source	1..1534				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
CDS	104..1258				
	/codon_start=1				
	/product="serine/threonine protein kinase"				
	/protein_id="CAE06249.1"				
	/db_xref="gi:32485022"				
	/translation="MTSPHPRLQDAAPKPKSRPNSDQFOEDLRQGFQKRLPGLASTYLNLKLGSGSYATYYKGISINQLVALYKSMNAEAGDPFIRASLLGKGLKMANIVLHDIITHEKTELFVEVETHTDLAQYMSHPGSHPHNYRFLPMLLEGKLVNIHVLHRLDKPOMLISHLELKLADPGARASITSSQYSEVYTLWRPPLALGATYSSSELDIWGACITFIEMFGQPLPFGVNLILEQLKEIWEVLGVFTDVFHVKLKNYNEWEPFLPTPRSLHVMNRLRVEADLAAQMLKPFPRDVSQDEALVHDYVSALPQQLYQLPDESLFTVSGVRLKPEMCDLLASYQGHNPQFQSKW"				
ORIGIN					
Query Match	90.6%;	Score 1185.4;	DB 6;	Length 1534;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 1186;	Conservative	0;	Mismatches	1;	Indels 0;
					Gaps 0;
QY	122 AGCTACAGACCCCAAAGAAGCATCATGTTCCATGACTTCATTTGACCCACGGAGCTTC	181			
Db	72 ATTAAACAGACCTAAAGAAGCATCATGTTCCATGACTTCATTTGACCCACGGAGCTTC	131			
QY	182 AACCTGCCCCGTCAGAGATTCAAGATGTAAGGCGACAGGATGAACATGATGTTTTTC	241			
Db	132 AACCTGCCCCGTCAGAGATTCAAGATGTAAGGCGACAGGATGAACATGATGTTTTTC	191			
QY	242 AGGAAGAGATCTGAGGACAGGAGTTTTCAGTGAAGGAAGAGCTCCCTTTTGGGGACGCT	301			
Db	192 AGGAAGAGATCTGAGGACAGGAGTTTTCAGTGAAGGAAGAGCTCCCTTTTGGGGACGCT	251			
QY	302 CATCTTACTTGAACCTTGAGAGAGCTGGGTGAAGGCTCTTATGGGACAGTTTCAAGAGGGA	361			
Db	252 CATCTTACTTGAACCTTGAGAGAGCTGGGTGAAGGCTCTTATGGGACAGTTTCAAGAGGGA	311			
QY	362 TTGAGCAATAATATGACACATAGTGCTTAAAGTCATCAAGCATGAATGACAGAGAAAG	421			
Db	312 TTGAGCAATAATATGACACATAGTGCTTAAAGTCATCAAGCATGAATGACAGAGAAAG	371			
QY	422 GAGTCCATTTACAGCTATCCGAGAGAGCTTCTCTCGAAGGGTTGAAACATGCCAATA	481			
Db	372 GAGTCCATTTACAGCTATCCGAGAGAGCTTCTCTCGAAGGGTTGAAACATGCCAATA	431			
QY	482 TTGGTCTCCGACAGACATATCCACACCAAGAGACATGACATTCGTTTTGAATACA	541			
Db	432 TTGGTCTCCGACAGACATATCCACACCAAGAGACATGACATTCGTTTTGAATACA	491			
QY	542 TGACACAGACCTGGCCAGTATATGTCTCAGCATCCAGAGGGGCTTCATCTCATATAG	601			
Db	492 TGACACAGACCTGGCCAGTATATGTCTCAGCATCCAGAGGGGCTTCATCTCATATAG	551			
QY	602 TGAACCTTATATGTTTAACTTTTGCGGGGCTCGGGGATACATCAACCAACACGTTTC	661			
Db	552 TGAACCTTATATGTTTAACTTTTGCGGGGCTCGGGGATACATCAACCAACACGTTTC	611			

Qy	662	TTTCACAGGAGACCTGAAACCTCAGAACTTATCTCATATGACACCTGGGAGAGCTCAAACTCG	721
Db	612	TTTCACAGGAGACCTGAAACCTCAGAACTTATCTCATATGACACCTGGGAGAGCTCAAACTCG	671
Qy	722	CTGATTTTGGTCTTGGCCGGGCGCAAGTCATCTCCAGCGCAGACATATCTTTCAAGAGTCG	781
Db	672	CTGATTTTGGTCTTGGCCGGGCGCAAGTCATCTCCAGCGCAGACATATCTTTCAAGAGTCG	731
Qy	782	TGACCCCTCTGGTACCGGACCCCTGATGCTTTGCTGGAGGCCATTAATAATTCCTGTAGC	841
Db	732	TGACCCCTCTGGTACCGGACCCCTGATGCTTTGCTGGAGGCCATTAATAATTCCTGTAGC	791
Qy	842	TGACAATAAGGGGGTGGAGGGCTGATCTTATTTGAAGATGTTCCAGGGTCACACTTTGTTTC	901
Db	792	TGACAATAAGGGGGTGGAGGGCTGATCTTATTTGAAGATGTTCCAGGGTCACACTTTGTTTC	851
Qy	902	CTGGGGGTTTTCCAACTCTTTGAAACAGCTGAGAAAATCTGGAGGTTGCTGGAGTCCCTTA	961
Db	852	CTGGGGGTTTTCCAACTCTTTGAAACAGCTGAGAAAATCTGGAGGTTGCTGGAGTCCCTTA	911
Qy	962	CAGAGATATCTTGGCCGGGAGTCTCCAACTACTCAATCAATCCAGAATGTTTCCGAC	1021
Db	912	CAGAGATATCTTGGCCGGGAGTCTCCAACTACTCAATCAATCCAGAATGTTTCCGAC	971
Qy	1022	TGCTTACGCTTCGAGAGCTTCATGTTGTCTGGAACAGGCTGGCAGGTTTCTGAAGCTG	1081
Db	972	TGCTTACGCTTCGAGAGCTTCATGTTGTCTGGAACAGGCTGGCAGGTTTCTGAAGCTG	1031
Qy	1082	AAGACTTGACCTCCCAATGCTTAAAGGCTTCCAGAGACGCGCTCCGCCCAGAGAG	1141
Db	1032	AAGACTTGACCTCCCAATGCTTAAAGGCTTCCAGAGACGCGCTCCGCCCAGAGAG	1091
Qy	1142	CACATGTTCAATGATTAATTTCAGCGCCCTGCACTCTCACTGTACCAAGCTTCTGATGAG	1201
Db	1092	CACATGTTCAATGATTAATTTCAGCGCCCTGCACTCTCACTGTACCAAGCTTCTGATGAG	1151
Qy	1202	AGCTTTTGTTTACAGTTTCAGAGATGAGGCTAAAGCCAGAAAATGTGTGACCTTTGGCGCT	1261
Db	1152	AGCTTTTGTTTACAGTTTCAGAGATGAGGCTAAAGCCAGAAAATGTGTGACCTTTGGCGCT	1211
Qy	1262	CCTACAGAAAGGTCACACACCAGCCAGCTTGTAGCAATGCTGGTGA	1308
Db	1212	CCTACAGAAAGGTCACACACCAGCCAGCTTGTAGCAATGCTGGTGA	1258
RESULT 9			
	AX772795	1534 bp	DNA linear PAT 09-JUN-2003
	LOCUS	AX772795	
	DEFINITION	Sequence 5 from Patent WO03046167.	
	ACCESSION	AX772795	
	VERSION	AX772795.1	GI:32485024
	KEYWORDS		
	SOURCE		
	ORGANISM	Homo sapiens (human)	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
	REFERENCE		
	1	Koehler, R. H.	
	AUTHORS	Regulation of human serine/threonine protein kinase	
	TITLE	Patent: WO 03046167-A 5 05-JUN-2003;	
	JOURNAL	Bayer Aktiengesellschaft (DE)	
	FEATURES	location/Qualifiers	
	source	1..1534	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
		/note="serine/threonine protein kinase"	
ORIGIN			
	Query Match	90 6%;	Score 1185.4; DB 6; Length 1534;
	Best Local Similarity	99.9%;	Pred. No. 0;
	Matches 1186; Conservative	0;	Mismatches 1; Indels 0; Gaps 0;

122 AGCTAACAGACCTAAAGAGCATGTCATGATCTTCAATTCAGCCAGGAGCTTC 181
 72 ATCTAACAGACCTAAAGAGCATGTCATGATCTTCAATTCAGCCAGGAGCTTC 131
 182 AAGCTGCGCGCCGAGAGTTCAGAGTAAAGGCGAGAGTAACAGTGAATGTTTTC 241
 132 AAGCTGCGCGCCGAGAGTTCAGAGTAAAGGCGAGAGTAACAGTGAATGTTTTC 191
 242 AGAAGAGAGATCTGAGGAGGCTTTCACTGAGAGAGAGCTCCCTTTTGGGAGCGCT 301
 132 AGAAGAGAGATCTGAGGAGGCTTTCACTGAGAGAGAGCTCCCTTTTGGGAGCGCT 251
 302 CATCTTACTTGAATCTTGAGAGAGCTGGGAGAGCTTTATGCGACAGTTTACAGAGGA 361
 252 CATCTTACTTGAATCTTGAGAGAGCTGGGAGAGCTTTATGCGACAGTTTACAGAGGA 311
 362 TTAGAGAGATTAATGAGACAACTAGTGGCTTTAAAGATCATGACATGAATGAGAGAG 421
 312 TTAGAGAGATTAATGAGACAACTAGTGGCTTTAAAGATCATGACATGAATGAGAGAG 371
 422 GAGTCCCATTTACAGCTATCCGAGAAAGCTTCTCTCGAAGGGTTGAAACATGCCAATA 481
 372 GAGTCCCATTTACAGCTATCCGAGAAAGCTTCTCTCGAAGGGTTGAAACATGCCAATA 431
 482 TTGTGCTCTGATGATGATATTCACACCAAGAGACATGACATTCGTTTGAATACA 541
 432 TTGTGCTCTGATGATGATATTCACACCAAGAGACATGACATTCGTTTGAATACA 491
 542 TGACACACAGACCTGGGCGGAGTATATGTCAGCATCCAGAGAGGCTTCATCTCATATATG 601
 492 TGACACACAGACCTGGGCGGAGTATATGTCAGCATCCAGAGAGGCTTCATCTCATATATG 551
 602 TCAGACTTTTCATGTTTCACTTTTGGGCGGCTGGCGTATCCACCAACACAGCTTC 661
 552 TCAGACTTTTCATGTTTCACTTTTGGGCGGCTGGCGTATCCACCAACACAGCTTC 611
 662 TTCAACAGGAGCTGAAACCTCAAGACTTATCTCATGATCACTGGGAGAGCTCAAACTGG 721
 612 TTCAACAGGAGCTGAAACCTCAAGACTTATCTCATGATCACTGGGAGAGCTCAAACTGG 671
 722 CTGATTTTGGTCTGTCGGGCGGAGTGCATTCCTCCAGCAGACATATCTCTTCAAGAGTGG 781
 672 CTGATTTTGGTCTGTCGGGCGGAGTGCATTCCTCCAGCAGACATATCTCTTCAAGAGTGG 731
 782 TGACCTCTGATACCGGCGGCGGCTGATGCTTGTGCGGAGCCACTGAATATTCCTCTGAGC 841
 732 TGACCTCTGATACCGGCGGCGGCTGATGCTTGTGCGGAGCCACTGAATATTCCTCTGAGC 791
 842 TGACATATGGGGTGGAGGCTGATCTTTATGAAATGTTCCAGGGTCAACTTTGTTTC 901
 792 TGACATATGGGGTGGAGGCTGATCTTTATGAAATGTTCCAGGGTCAACTTTGTTTC 851
 902 CTGGGGGTTTCAACATCTCTTGAACAGCTGAGAGAAATCTGGGAGGCTGGGAGTCCCTTA 961
 852 CTGGGGGTTTCAACATCTCTTGAACAGCTGAGAGAAATCTGGGAGGCTGGGAGTCCCTTA 911
 962 CAGAGGATCTTGGCGGAGGCTCTCCAGACTACTTAATCAATCCAGATGGTTCCTCCAC 1021
 912 CAGAGGATCTTGGCGGAGGCTCTCCAGACTACTTAATCAATCCAGATGGTTCCTCCAC 971
 1022 TGCCTACGCGCTGAAAGCTTCAATGTTGTGGAACAGGCTGGGAGGAGGTTCTGAAAGCTG 1081
 972 TGCCTACGCGCTGAAAGCTTCAATGTTGTGGAACAGGCTGGGAGGAGGTTCTGAAAGCTG 1031
 1082 AAGACCTGGCTCCCGAGATGCTAAAGAGGCTTCCCGAGAGACGCGCTCCGCGCAGAGAG 1141
 1032 AAGACCTGGCTCCCGAGATGCTAAAGAGGCTTCCCGAGAGACGCGCTCCGCGCAGAGAG 1091
 1142 CACTTGTTCATGATATTTCAAGGCGCTGCAATCTCAGCTGTATCCAGCTTCTGATAGAG 1201
 1092 CACTTGTTCATGATATTTCAAGGCGCTGCAATCTCAGCTGTATCCAGCTTCTGATAGAG 1151
 1202 AGCTTTTGTTCATGATATTTCAAGAGTGAAGGCTAAAGCAGAAATGTGTGACTTTTGGCCT 1261

Db 1152 AGCTTTTGTTCATGATATTTCAAGAGTGAAGGCTAAAGCAGAAATGTGTGACTTTTGGCCT 1211
 Qy 1262 CCTACAGAAAGATGACCAACCAGCCAGATTAGCAAAATGCTGTGA 1308
 Db 1212 CCTACAGAAAGATGACCAACCAGCCAGATTAGCAAAATGCTGTGA 1258
 RESULT 10
 AB053308 1534 bp mRNA linear PRI 05-OCT-2001
 LOCUS Homo sapiens AL52CR7 mRNA, complete cds..
 DEFINITION
 ACCESSION AB053308
 VERSION AB053308.1 GI:15823641
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS
 1 Hadano, S., Hand, C.K., Ouga, H., Yanagisawa, Y., Ootomo, A.,
 Devon, R.S., Miyamoto, N., Showguchi-Miyata, J., Okada, Y.,
 Singaraja, R., Figlewicz, D.A., Kwiatkowski, T., Hosler, B.A.,
 Sasie, T., Skaug, J., Nasir, J., Brown, R.H. Jr., Scherer, S.W.,
 Rouleau, G.A., Hayden, M.R. and Ikeda, J.-E.
 TITLE A gene encoding a putative GTPase regulator is mutated in familial
 amyotrophic lateral sclerosis 2
 JOURNAL Nat. Genet. 29 (2), 166-173 (2001)
 MEDLINE 21470351
 PUBMED 11586298
 REFERENCE
 AUTHORS
 2 (bases 1 to 1534)
 Hadano, S. and Ikeda, J.
 TITLE Direct Submission
 JOURNAL Submitted (08-JAN-2001) Shinji Hadano, Tokai University, The
 Institute of Medical Sciences; Bohseida, Isehara, Kanagawa
 259-1193, Japan (E-mail: shinji@nga.med.u-tokai.ac.jp,
 Tel:81-463-91-5095, Fax:81-463-91-4993)
 FEATURES
 source
 1. 1534
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2q33"
 /clone="CALS124"
 /tissue_type="testis"
 /note="RT-PCR
 product:1-762nt-clone;IMAGE:730574;670-1534nt-composite
 sequence:1-1534nt"
 1. 1534
 /gene="AL52CR7"
 104. 1258
 /gene="AL52CR7"
 /codon_start=1
 /protein_id="BAB69017.1"
 /db_xref="GI:15823642"
 /translation="MTSPHPRGLQARAQKFKSRKPRNSDFQEBDLRQGFQWRKSL
 PFGASVYLNLEKLGSGSVATVYKGISRINQVLVLYISMAEAGVFTAIRASLL
 KGLKHANVILHDIHTKETLTFEVEVMTDLAQMSPGSLHPHNVLRFMFOILRG
 LAYITHQVLRDLKRPQNLISHLGELILARGLARARSIPQSYSEVVTLVWVPPD
 ALGATREYSSELDITWAGCITFEMEGQPLRPFGVSNIIIEQLKIEVLYGVPEDTWPG
 VSKLPNVNPEWPLPTPRSLHVNVRLLGRVPEADLASQMLKGFPRDRSAQELVHD
 YFSALPQSLYQLPDBESLFTVSGVRLKPEMDLLASLYQGHHPAQFSKMW"

ORIGIN

Query Match 90.6%; Score 1185.4; DB 9; Length 1534;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 122 AGCTAACAGACCTAAAGAGCATGTTCCATGATCTTCAATTCAGCCAGGAGCTTC 181
 Db 72 ATCTAACAGACCTAAAGAGCATGTTCCATGATCTTCAATTCAGCCAGGAGCTTC 131

QY 182 AAGCTGCGCGTCCGAGAGTTCAAGATGTAAGGCCACGAGTAAACAGTATGTTTTC 241
DB 132 AAGCTGCGCGTCCGAGAGTTCAAGATGTAAGGCCACGAGTAAACAGTATGTTTTC 191
QY 242 AAGAAAGAGATCTGAGAGCGGGGTTTCACTGAGAGAGAGCCCTCTTTTGGGGGAGCCCT 301
DB 192 AAGAAAGAGATCTGAGAGCGGGGTTTCACTGAGAGAGAGCCCTCTTTTGGGGGAGCCCT 251
QY 302 CATCTTACTTGAACCTTGAGAAAGCTGGGTGAAGGCTCTTATGCGACAGTTTTCAGAGGGA 361
DB 252 CATCTTACTTGAACCTTGAGAAAGCTGGGTGAAGGCTCTTATGCGACAGTTTTCAGAGGGA 311
QY 362 TTAGCGAATAATGAGCACTATGAGCTTTTAAAGTCAATCAGATGAATGCGAGAGAG 421
DB 312 TTAGCGAATAATGAGCACTATGAGCTTTTAAAGTCAATCAGATGAATGCGAGAGAG 371
QY 422 GAGTCCCATTTACAGATCTCCGAGAGGCTCTCTCCGTAAGGGGTTTGAACATGCCAATA 481
DB 372 GAGTCCCATTTACAGATCTCCGAGAGGCTCTCTCCGTAAGGGGTTTGAACATGCCAATA 431
QY 482 TTGTGCTCCTGCATGACATATATCCACACCAAGAGACACTGACATTCGTTTTGAATACA 541
DB 432 TTGTGCTCCTGCATGACATATATCCACACCAAGAGACACTGACATTCGTTTTGAATACA 491
QY 542 TGGACACAGACCTGGCGCCAGATATATGCTTCAGACATCCAGAGGGGCTTCACTCATATAG 601
DB 492 TGGACACAGACCTGGCGCCAGATATATGCTTCAGACATCCAGAGGGGCTTCACTCATATAG 551
QY 602 TCAGACCTTTTCAATGTTTCAACTTTTGGGGGCTGGGGTATCATCCACCAACAGCTTC 661
DB 552 TCAGACCTTTTCAATGTTTCAACTTTTGGGGGCTGGGGTATCATCCACCAACAGCTTC 611
QY 662 TTCACAGGAGCCTGAAACCTCAGAACTTATCTCATGATCACTGGAGAGAGCTTCAACTGG 721
DB 612 TTCACAGGAGCCTGAAACCTCAGAACTTATCTCATGATCACTGGAGAGAGCTTCAACTGG 671
QY 722 CTGATTTTGTCTTGGCGGGGCGAAGTCCATTTCCAGCCAGCAATACTCTTCAGAAAGTG 781
DB 672 CTGATTTTGTCTTGGCGGGGCGAAGTCCATTTCCAGCCAGCAATACTCTTCAGAAAGTG 731
QY 782 TGAACCTCTGATGACCGGGGCTGATGCTTTGCTGGGAGGCACTGATATTTCTCTGAGC 841
DB 732 TGAACCTCTGATGACCGGGGCTGATGCTTTGCTGGGAGGCACTGATATTTCTCTGAGC 791
QY 842 TGGACATATGAGGAGTGAAGGCTGATCTTTTATGAAATGTTTCAAGGTCACCTTTGTTTC 901
DB 792 TGGACATATGAGGAGTGAAGGCTGATCTTTTATGAAATGTTTCAAGGTCACCTTTGTTTC 851
QY 902 CTGGGGGTTTCCACATCTTTGAAACAGCTGAGAAAAATCTGGGAGAGTGTGGAGTCCCTA 961
DB 852 CTGGGGGTTTCCACATCTTTGAAACAGCTGAGAAAAATCTGGGAGAGTGTGGAGTCCCTA 911
QY 962 CAGAGGATACCTTGGCGGGAGTCTCCAAAGTACCTTAACTAATCCAGAAATGTTTCCAC 1021
DB 912 CAGAGGATACCTTGGCGGGAGTCTCCAAAGTACCTTAACTAATCCAGAAATGTTTCCAC 971
QY 1022 TGGCTTACGCTCGAAGGCTTCAATGTTGTGAAACAGGCTGGGACAGGTTTCTGAAAGCTG 1081
DB 972 TGGCTTACGCTCGAAGGCTTCAATGTTGTGAAACAGGCTGGGACAGGTTTCTGAAAGCTG 1031
QY 1082 AAAAACTGGGCTCCAGATGCTTAAAGGCTTTTCCAGAGACCGGCTCTCGGCCAGAGAG 1141
DB 1032 AAAAACTGGGCTCCAGATGCTTAAAGGCTTTTCCAGAGACCGGCTCTCGGCCAGAGAG 1091
QY 1142 CACTTGTTCATGATTTTCAAGGAGGCTTTCAGGAGGCTTTCAGCTGACAGCTTCTGATAGG 1201
DB 1092 CACTTGTTCATGATTTTCAAGGAGGCTTTCAGGAGGCTTTCAGCTGACAGCTTCTGATAGG 1151
QY 1202 AGCTTTTGTTCATGATTTTCAAGAGTGAAGGCTTAAAGCCAGAAATGTGTGACCTTTTGGCCT 1261
DB 1152 AGCTTTTGTTCATGATTTTCAAGAGTGAAGGCTTAAAGCCAGAAATGTGTGACCTTTTGGCCT 1211
QY 1262 CCAACCAAGAAAGGTCAACCAAGCCCAAGTTTACCAAAATGTGTGTA 1308

DB 1212 CCTACCAAGAAAGGTCAACCAAGCCCAAGTTTACCAAAATGTGTGTA 1258
RESULT 11
AR265350
LOCUS AR265350 2203 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6492154.
ACCESSION AR265350
VERSION AR265350.1 GI:2963853
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .2203
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 77.1%; Score 1009; DB 6; Length 2203;
Best Local Similarity 100.0%; Pred. No. 4,6e-313; Mismatches 0; Gaps 0;
Matches 1009; Conservative 0; Indels 0; Gaps 0;
QY 1 AAGGGGTCAAGAGCTGTGTCGCAAGACTGATACAGCCCTGATGAGCTGATCACTTGTTC 60
DB 118 ATGGGTCAAGAGCTGTGTCGCAAGACTGATACAGCCCTGATGAGCTGATCACTTGTTC 177
QY 61 GAGGAGGCGGAGGACACAGCTGTGAGAGACTGAGCTGAGACACGAGGCTGCGTTC 120
DB 178 GAGGAGGCGGAGGACACAGCTGTGAGAGACTGAGCTGAGACACGAGGCTGCGTTC 237
QY 121 AAGCTTACAGACCTTAAAGACATCATGTTTCAATGACTTTCATTTTCAACCCAGGGACTT 180
DB 238 AAGCTTACAGACCTTAAAGACATCATGTTTCAATGACTTTCATTTTCAACCCAGGGACTT 297
QY 181 CAAGCTGCGCGTCCGAGAGGTTCAAGAGTAAAGGCCACGAGTAAACAGTATGTTTTC 240
DB 298 CAAGCTGCGCGTCCGAGAGGTTCAAGAGTAAAGGCCACGAGTAAACAGTATGTTTTC 357
QY 241 CAGAAAGAGATCTGAGCAGGAGTTTCAAGTGAAGAAAGGCTCTTTTGGGGCAGGC 300
DB 358 CAGAAAGAGATCTGAGCAGGAGTTTCAAGTGAAGAAAGGCTCTTTTGGGGCAGGC 417
QY 301 TCATCTTACTTGAACCTTGAAGAAAGCTGGGTGAAGGCTCTTATGCCACAGTTTACAAGGGG 360
DB 418 TCATCTTACTTGAACCTTGAAGAAAGCTGGGTGAAGGCTCTTATGCCACAGTTTACAAGGGG 477
QY 361 ATTAGCAGAAATAATGAGCACTAGTGGCTTTTAAAGTCAATCAGATGAATGCAAGAGAA 420
DB 478 ATTAGCAGAAATAATGAGCACTAGTGGCTTTTAAAGTCAATCAGATGAATGCAAGAGAA 537
QY 421 GGAATCCATTTACAGCTATCCGAGAAAGCTTCTCTGAAAGGTTTGAACATGCAAT 480
DB 538 GGAATCCATTTACAGCTATCCGAGAAAGCTTCTCTGAAAGGTTTGAACATGCAAT 597
QY 481 ATTGTGCTCTGTCAGATCAATATCCACACCAAGAGACCTGACATTCCTTTTGAATAC 540
DB 598 ATTGTGCTCTGTCAGATCAATATCCACACCAAGAGACCTGACATTCCTTTTGAATAC 657
QY 541 ATGACACAGACCTGGCCAGATATATGCTCAGATATCCAGAGGAGGCTTCAATCTCAATAT 600
DB 658 ATGACACAGACCTGGCCAGATATATGCTCAGATATCCAGAGGAGGCTTCAATCTCAATAT 717
QY 601 GTTCAGACTTTTCAATGTTTCACTTTTGGGGGCTGGCGTATCATCCACCAACAGCTT 660
DB 718 GTTCAGACTTTTCAATGTTTCACTTTTGGGGGCTGGCGTATCATCCACCAACAGCTT 777

DB 358 CAGGAGAGGATCTGAGGCAAGGTTTCAGTGG

```
/db_xref="taxon:  
/note="Incyte
```

ORIGIN	
Query Match	67.9%; Score 887.6; DB 6; Length 1376;
Best Local Similarity	99.6%; Pred. No. 6.3e-274;
Matches	890; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	122 AGCTAACAGACCTTAAGAAAGCATCATGTTCCATGATCTTCACTTTCACCCAGGGGACTTC 181
DB	266 ATCTTAACAGACCTTAAGAAAGCATCATGTTCCATGATCTTCACTTTCACCCAGGGGACTTC 325
QY	182 AACCTGCCCGCCGAGAGTTCAAGAGTAAAGGCCAGAGTAACAGATGTTGTTTC 241
DB	326 AACCTGCCCGCCGAGAGTTCAAGAGTAAAGGCCAGAGTAACAGATGTTGTTTC 385
QY	242 AGGAAGAGATCTGAGCAGAGGTTTTCAGTGAAGAAAGCTCCCTTTTGGGGAGCCCT 301
DB	386 AGGAAGAGATCTGAGCAGAGGTTTTCAGTGAAGAAAGCTCCCTTTTGGGGAGCCCT 445
QY	302 CATCTTAAGATCTGAGAGAAAGCTGGTGAAGCTTTATGGGACAGTTTCAAGGGGA 361
DB	446 CATCTTAAGATCTGAGAGAAAGCTGGTGAAGCTTTATGGGACAGTTTCAAGGGGA 505
QY	362 TTGAGCAATTAATGAGCACTAGTGGCTTTAAAGTCAATCAGATGAATGACAGAGAA 421
DB	506 TTGAGCAATTAATGAGCACTAGTGGCTTTAAAGTCAATCAGATGAATGACAGAGAA 565
QY	422 GAGTCCCATTTACAGATCTCCGAGAAAGCTTCTCTCTGAAGGGTTGAAACATGCCAAT 481
DB	566 GAGTCCCATTTACAGATCTCCGAGAAAGCTTCTCTCTGAAGGGTTGAAACATGCCAAT 625
QY	482 TTGTCCTCCCGCAGCATATATCCACCAAGACATGACATTCGTTTGAATPAC 541
DB	626 TTGTCCTCCCGCAGCATATATCCACCAAGACATGACATTCGTTTGAATPAC 685
QY	542 TGCAACAGAGCTGGCCAGATATATGTCAGCATCCAGAGGGCTTCATCATATATG 601
DB	686 TGCAACAGAGCTGGCCAGATATATGTCAGCATCCAGAGGGCTTCATCATATATG 745
QY	602 TCAAGCTTTTCATGTTTCAACTTTTGGGGGGCTGGGCTACATCCACACCAACAGGTT 661
DB	746 TCAAGCTTTTCATGTTTCAACTTTTGGGGGGCTGGGCTACATCCACACCAACAGGTT 805
QY	662 TTCAAGGGGAGCTGGAACCTCAGAACTTCACTCATGTCACCTGGGAGAGCTCAAACTGG 721
DB	806 TTCAAGGGGAGCTGGAACCTCAGAACTTCACTCATGTCACCTGGGAGAGCTCAAACTGG 865
QY	722 CTGATTTTGGTCTTGGCCCGGAGTCAATCCAGCCAGACATATCTTTCAAGAGTGG 781
DB	866 CTGATTTTGGTCTTGGCCCGGAGTCAATCCAGCCAGACATATCTTTCAAGAGTGG 925
QY	782 TGAACCTCTGTACCGGCCCCCTGTATGCTTGGGAGCCACTGAATATTTCTGTAGC 841
DB	926 TGAACCTCTGTATCCGGCCCCCTGTATGCTTGGGAGCCACTGAATATTTCTGTAGC 985
QY	842 TGAACATATGGGGTGAAGGCTGATCTTTATGAAATGTTCAAGGGTCAACTTTGTTTC 901
DB	986 TGAACATATGGGGTGAAGGCTGATCTTTATGAAATGTTCAAGGGTCAACTTTGTTTC 1045
QY	902 CTGGGGTTTCCACATCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCTTA 961
DB	1046 CTGGGGTTTCCACATCTTGAACAGCTGAGAAATCTGGAGGTGCTGGAGTCCCTTA 1105
QY	962 CAGAGGATCTTGGCCGGAGTCTCCAAAGTACCTAATCAATCCGAATGGT 1015
DB	1106 CAGAGGATCTTGGCCGGAGTCTCCAAAGTACCTAATCAATCCGAAGAGT 1159
RESULT 14	
LOCUS	AK131512 1828 bp mRNA linear PRI 07-MAY-2004
DEFINITION	Homo sapiens cDNA FLJ16732 fis, clone BNGH42005017, moderately similar to SERINE/THREONINE-PROTEIN KINASE PCTAIRE-1 (EC 2.7.1.-).
ACCESSION	AK131512

VERSION	AK131512.1 GI:47077537
KEYWORDS	oligo capping, fis (full insert sequence).
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Makamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Magatsuma, M., Murakawa, K., Kanehori, K., Takahashi, Fujii, A., Oshima, A., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE	NEDO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1828)
AUTHORS	Isogai, T. and Yamamoto, J.
JOURNAL	Submitted (01-MAR-2004) Takao Isogai, Fuj Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kitarazu, Chiba 292-0818, Japan Fax: 81-438-52-1986 (E-mail: genom@shri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB) (supported by Japan construction: Heijix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES	Location/Qualifiers
source	1..1828
	/organism="Homo sapiens"
	/mol_type="RNA"
	/db_xref="taxon:9606"
	/clone="BNGH42005017"
	/cell_line="H4"
	/cell_type="neurogloma"
	/clone_lib="BNGH42"
	/note="cloning vector: pME18P13"
CDS	87..1124
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="BAD18656.1"
	/db_xref="GI:47077538"
	/translation="MGDELCAKTYPGGSCYHSCSGRAGHGRSOPETTRAFKLTDLKASMSFPHRGLQARAKQKSKPRNSDGFQEDLRQGMKSKPRGAASTYLNLRKLGESYATYKGISRLNQLVALKYISMAEGVPEPTAIRLSLKLGHANLYLHLDIIHTKELTFVEFYMLFMQLRGALYIHQVLIHNRDLKPNMLLSHGLKLDLFGARAKSIPQSYSEVVTLMYRPDNLGATSEYSELIDWGACGPIEMEQOGLPFGVSNILRQLKIMWEVLGVPTEBDTPGVSKLPNYPGSKKQGHMHTDIGNRTAAMQPSRNVSF"
ORIGIN	
Query Match	66.7%; Score 873; DB 9; Length 1828;
Best Local Similarity	93.8%; Pred. No. 3.3e-269;
Matches	946; Conservative 0; Mismatches 0; Indels 63; Gaps 1;
QY	1 ATGGGTCAAGAGCTGTGTGCAAGAGCTGTACAGCTGTGATGAGCTGCTACCATTTGTTCA 60
DB	87 ATGGGTCAAGAGCTGTGTGCAAGAGCTGTACAGCTGTGATGAGCTGCTACCATTTGTTCA 146
QY	61 GAGGAGGAGGAGGACACAGCTGTGGAGAGAGTCAAGCTGAGACCAAGGAGGCTCGTTTC 120
DB	147 GAGGAGGAGGAGGACACAGCTGTGGAGAGAGTCAAGCTGAGACCAAGGAGGCTCGTTTC 206
QY	121 AAGCTAACAGACTTAAGAAAGCATCATGTTTCATGATCTTCACTTTCACCCAGGGGACTT 180
DB	207 AAGCTAACAGACTTAAGAAAGCATCATGTTTCATGATCTTCACTTTCACCCAGGGGACTT 266
QY	181 CAAAGTCCCGGCGCCAGAGTTTCAAGAGTAAAGGCCAGAGTAAACAGTAACTGTTT 240
DB	267 CAAAGTCCCGGCGCCAGAGTAAAGGCCAGAGTAAACAGTAACTGTTT 326


```

PT /number= 3
PT replace(3673,G)
PT /tag= k
PT variation
PT replace(3747,G)
PT /tag= l
PT variation
PT replace(3788,G)
PT /tag= m
PT exon
PT 7939. .8018
PT /tag= n
PT /number= 4
PT intron
PT 8019. .18260
PT /tag= o
PT /number= 4
PT variation
PT replace(8034,T)
PT /tag= p
PT exon
PT 18261. .18355
PT /tag= q
PT /number= 5
PT intron
PT 18356. .19116
PT /tag= r
PT /number= 5
PT exon
PT 19117. .19179
PT /tag= s
PT /number= 6
PT intron
PT 19180. .29296
PT /tag= t
PT /number= 6
PT replace(27740,C)
PT /tag= u
PT variation
PT replace(27752,T)
PT /tag= v
PT exon
PT 29297. .29420
PT /tag= w
PT /number= 7
PT intron
PT 29421. .31091
PT /tag= x
PT /number= 7
PT variation
PT replace(29927,T)
PT /tag= y
PT variation
PT replace(30772,G)
PT /tag= z
PT exon
PT 31092. .31212
PT /tag= aa
PT /number= 8
PT intron
PT 31213. .42825
PT /tag= ab
PT /number= 8
PT variation
PT replace(36310,G)
PT /tag= ac
PT variation
PT replace(36327,T)
PT /tag= ad
PT variation
PT replace(40618,C)
PT /tag= ae
PT variation
PT replace(40928,T)
PT /tag= af
PT exon
PT 42826. .42919
PT /tag= ag
PT /number= 9
PT intron
PT 42920. .50322
PT /tag= ah
PT /number= 9
PT variation
PT replace(44701,A)
PT /tag= aj
PT variation
PT replace(46020,C)
PT /tag= ak
PT variation
PT replace(46036,G)
PT /tag= al
PT variation
PT replace(46095,A)
PT /tag= am
PT variation
PT replace(47608,G)
PT /tag= an
PT exon
PT 50323. .50409
PT /tag= ao

```

```

FT /number= 10
FT variation replace(52150,A)
FT /tag= ap
XX
XX WO200261060-A2.
XX
XX
XX 08-AUG-2002.
XX
XX PF
XX 17-JAN-2002; 2002WO-US001106.
XX
XX PR
XX 31-JAN-2001; 2001US-0265151P.
XX PR
XX 09-MAR-2001; 2001US-00801861.
XX
XX (PEKE ) PE CORP NY.
XX
XX PI
XX Yan C, Ketchum K, Di Francesco V, Beasley EM;
XX
XX MPI; 2002-608515/65.
XX DR
XX P-PSDB; AAO18613, AAO18614.
XX
XX PS
XX Claim 4; Fig 3; 131pp; English.
XX
XX CC The present invention provides the protein, cDNA and gene sequences of
XX two splice variants of a human pftalre family kinase. The sequences are
XX specifically expressed in the human testis, brain, uterus endometrium
XX adenocarcinomas, lung fibroblasts, kidney renal cell adenocarcinomas, and
XX CC can be used to treat related diseases. The present sequence is the gene
XX of the invention
XX
SQ Sequence 53332 BP; 15976 A; 10335 C; 10952 G; 16069 T; 0 U; 0 Other;

Query Match 100.0%; Score 5000; DB 6; Length 53332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TATAGGCCAATGCTGTGCTCAGCGGTATTTCCAGACCTTTGGAGGACAGAGATCG 60
DB 1 TATAGGCCAATGCTGTGCTCAGCGGTATTTCCAGACCTTTGGAGGACAGAGATCG 60
QY 61 CTTGAGCTCAGGAATTTGAGAGACAGACCTACGTAATAGTGAACCTCTGTCTGTACAA 120
DB 61 CTTGAGCTCAGGAATTTGAGAGACAGACCTACGTAATAGTGAACCTCTGTCTGTACAA 120
QY 121 TATATAAAGAAATTTTCCAGGACATGGTGGCGTSCACCCAGTGCAGCTATTGGAGGC 180
DB 121 TATATAAAGAAATTTTCCAGGACATGGTGGCGTSCACCCAGTGCAGCTATTGGAGGC 180
QY 181 TGAGGTAGGAGAAATGCTTGAAGCCAGAGGTTGAAAGACAGCCTTAGCAATAGTGAGA 240
DB 181 TGAGGTAGGAGAAATGCTTGAAGCCAGAGGTTGAAAGACAGCCTTAGCAATAGTGAGA 240
QY 241 CCTGTGTATATAAATAATTAAGTGTCTTCTTGACAGCGCTCGAGCTAGTAAGT 300
DB 241 CCTGTGTATATAAATAATTAAGTGTCTTCTTGACAGCGCTCGAGCTAGTAAGT 300
QY 241 CCTGTGTATATAAATAATTAAGTGTCTTCTTGACAGCGCTCGAGCTAGTAAGT 300
DB 241 CCTGTGTATATAAATAATTAAGTGTCTTCTTGACAGCGCTCGAGCTAGTAAGT 300
QY 301 CGAAGACTGAGGTGGAGATCATCTGAGCCAGAGAGGCTGAGCTGAGTAACAGTGA 360
DB 301 CGAAGACTGAGGTGGAGATCATCTGAGCCAGAGAGGCTGAGCTGAGTAACAGTGA 360
QY 361 TCACCCAGCTGATTCGAGCCTCGAAGACAGAGGGAGACCTGTTCCAAAAA 420
DB 361 TCACCCAGCTGATTCGAGCCTCGAAGACAGAGGGAGACCTGTTCCAAAAA 420
QY 421 AAAAAAAAAATGCAAGAAAAAGATCATTAACCTTGACCTGGGACATTAACCTTTATGTGATG 480
DB 421 AAAAAAAAAATGCAAGAAAAAGATCATTAACCTTGACCTGGGACATTAACCTTTATGTGATG 480
QY 481 AATTACAAATCTTTAGGAGAAATTAGCATTTCTGATTAATGATATATATT 540
DB 481 AATTACAAATCTTTAGGAGAAATTAGCATTTCTGATTAATGATATATATT 540

```


Db	481	AAATTCACAACTCTTTAGAGAAATTAGCAATTTCTGATAAATGATTAATTAATTT	540
Qy	541	ATTATTAATTCAAATGGAATTAAATTTCTGAGAACTAGCTTCTCACTCTCAGTGT	600
Db	541	ATTATTAATTCAAATGGAATTAAATTTCTGAGAACTAGCTTCTCACTCTCAGTGT	600
Qy	601	CAGTCAAAACCTTTAATGCTCTTTGGCCGGGTCCGTGCTCAGCCTGTATCCAGAC	660
Db	601	CAGTCAAAACCTTTAATGCTCTTTGGCCGGGTCCGTGCTCAGCCTGTATCCAGAC	660
Qy	661	TTTGGAGGCCAGAGCGGGTGGATCACAAGGTTAGAGATCGAGCAATCCGCGCTACA	720
Db	661	TTTGGAGGCCAGAGCGGGTGGATCACAAGGTTAGAGATCGAGCAATCCGCGCTACA	720
Qy	721	CGGTGAAAACCTGCTCTCTAATAAATACAAAAAATAGCCGGGTCGGTGCAGACGCC	780
Db	721	CGGTGAAAACCTGCTCTCTAATAAATACAAAAAATAGCCGGGTCGGTGCAGACGCC	780
Qy	781	TGTAGTCCCAAGCTGCTCAGAGAGCTGAGGCAGAGATGCTGTGAACCCGGAGCGAG	840
Db	781	TGTAGTCCCAAGCTGCTCAGAGAGCTGAGGCAGAGATGCTGTGAACCCGGAGCGAG	840
Qy	841	CTTGCAAGTACGCCAGATTGGCCCACTGCACTCCAGCCCTGGGCCGACATGGAGACTGT	900
Db	841	CTTGCAAGTACGCCAGATTGGCCCACTGCACTCCAGCCCTGGGCCGACATGGAGACTGT	900
Qy	901	TCTCAAAAAAATTTTTTTTTTTTTTTTCAAAAAATATCTAGATTGAATCTTGGAAATG	960
Db	901	TCTCAAAAAAATTTTTTTTTTTTTTTTCAAAAAATATCTAGATTGAATCTTGGAAATG	960
Qy	961	TTCTTCTTCTTTTTTTTTTTTTTTTCAAAAAATATCTAGATTGAATCTTGGAAATG	1020
Db	961	TTCTTCTTCTTTTTTTTTTTTTTTTCAAAAAATATCTAGATTGAATCTTGGAAATG	1020
Qy	1021	TAAAGTCTCTTCTCTTGTGGCAATTTTGAATGAAAAATACAGCTCATTAATTAATAC	1080
Db	1021	TAAAGTCTCTTCTCTTGTGGCAATTTTGAATGAAAAATACAGCTCATTAATTAATAC	1080
Qy	1081	CTGAACATTTTAAAAAAACCATCATGAGGTTCAAAATATCAAAATATTAATTAATGTTG	1140
Db	1081	CTGAACATTTTAAAAAAACCATCATGAGGTTCAAAATATCAAAATATTAATTAATGTTG	1140
Qy	1141	GATTAATAGACATPACTCTAATTTTTTCCCTTAATATGATGTTTATATATCTTCATTC	1200
Db	1141	GATTAATAGACATPACTCTAATTTTTTCCCTTAATATGATGTTTATATATCTTCATTC	1200
Qy	1201	TGCTTCACCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1260
Db	1201	TGCTTCACCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1260
Qy	1261	AGGTGTCATTTGAAGACTACAGACCTGCAAAACCAATTTAAGCAGATTCCTTGACATGT	1320
Db	1261	AGGTGTCATTTGAAGACTACAGACCTGCAAAACCAATTTAAGCAGATTCCTTGACATGT	1320
Qy	1321	GTGCTGTAAATATAGTACTTTACATATAGTAACATTAATTAATTAATTAATTAATTA	1380
Db	1321	GTGCTGTAAATATAGTACTTTACATATAGTAACATTAATTAATTAATTAATTAATTA	1380
Qy	1381	CAAGCAAGAAAGAAAGAAAGTATTTCAATCAAACTCTCTCTCCATCACCATTGGCTA	1440
Db	1381	CAAGCAAGAAAGAAAGAAAGTATTTCAATCAAACTCTCTCTCCATCACCATTGGCTA	1440
Qy	1441	ATATCATCTTTGTACAGTTAAGAAACAATATGAGTCTCACCAATATGTTTGAATTA	1500
Db	1441	ATATCATCTTTGTACAGTTAAGAAACAATATGAGTCTCACCAATATGTTTGAATTA	1500
Qy	1501	TGAATGAATGGCAACCTTCTTAAGACTAATGGAATACATATTGTTTGAAGCAAGAGAT	1560
Db	1501	TGAATGAATGGCAACCTTCTTAAGACTAATGGAATACATATTGTTTGAAGCAAGAGAT	1560
Qy	1561	GCAATGATATTTTCACTTTTCTCTGTTTATGATTCGTGGTTCTTTGACTACTAA	1620

Db	1561	GCAGTAGATATTTTCAACTTTTTTCGCTTTTAAGTATCTGCGTTCTTTGACGACGAA	1621
Qy	1621	AAGTAGTAGTAGAGCAAAATTTGTTTTAAAGCTCGAAAAACCAAAATGCTTTCAGATAAAA	1680
Db	1621	AAGTAGCTAGTAGCAAAATTTGTTTTAAAGCTCGAAAAACCAAAATGCTTTCAGATAAAA	1680
Qy	1661	GGTAGGGGAAAAAATCTCCCTCAACATGCTCACTTTAGCACACAGAAAAACCTAATATCAAA	1740
Db	1661	GGTAGGGGAAAAAATCTCCCTCAACATGCTCACTTTAGCACACAGAAAAACCTAATATCAAA	1740
Qy	1741	TATCACATCATGATGATCATATATAAATATCATTTGCAATAGATAAGCAATGTCATATCCCTAA	1800
Db	1741	TATCACATCATGATGATCATATATAAATATCATTTGCAATAGATAAGCAATGTCATATCCCTAA	1800
Qy	1801	AAACTATGTATACCAATAGCACTAACTGTGGCCAGAAACAAGAACTTTAACTGTGCCAAA	1860
Db	1801	AAACTATGTATACCAATAGCACTAACTGTGGCCAGAAACAAGAACTTTAACTGTGCCAAA	1860
Qy	1861	TTTTATTTCTATTCATATAACAGCTGCTCTGTTTTCAAGTTGTGCACATCTGAATGCAAGCA	1920
Db	1861	TTTTATTTCTATTCATATAACAGCTGCTCTGTTTTCAAGTTGTGCACATCTGAATGCAAGCA	1920
Qy	1921	TCCCTGCTGATGATGAGAGTTTCTTGCACTGATAAGAAAAAATCTGTAAGTTGTGAGGCT	1980
Db	1921	TCCCTGCTGATGATGAGAGTTTCTTGCACTGATAAGAAAAAATCTGTAAGTTGTGAGGCT	1980
Qy	1981	GCTCCAGGACAGAGCCATGATGATCATATGAAAAGCTCAACGCTGCTGACCTCTGGCAA	2040
Db	1981	GCTCCAGGACAGAGCCATGATGATCATATGAAAAGCTCAACGCTGCTGACCTCTGGCAA	2040
Qy	2041	AAAGGGAGAGAACAAAGATAGAGAGGCAAGTGGGGGAAAGTTCAAGTGCGGGTTTTCTC	2100
Db	2041	AAAGGGAGAGAACAAAGATAGAGAGGCAAGTGGGGGAAAGTTCAAGTGCGGGTTTTCTC	2100
Qy	2101	CTTGAACTTAACAAGTTATGGGTGAAGAGCTGTGTGCAAAACCTGTCAGCTGTGAGTCA	2160
Db	2101	CTTGAACTTAACAAGTTATGGGTGAAGAGCTGTGTGCAAAACCTGTCAGCTGTGAGTCA	2160
Qy	2161	GCTGCTACCATTTGTTCAAGAGGAGGCGAGGACACAGCTGTGCGAGAGAGTCAAGCTGAGA	2220
Db	2161	GCTGCTACCATTTGTTCAAGAGGAGGCGAGGACACAGCTGTGCGAGAGAGTCAAGCTGAGA	2220
Qy	2221	CCAGCGAGGCTGCGTTCAAGGTATTTGTATCCACAGAGAGACATCTTTCTCTATTTGATA	2280
Db	2221	CCAGCGAGGCTGCGTTCAAGGTATTTGTATCCACAGAGAGACATCTTTCTCTATTTGATA	2280
Qy	2281	AAACCAAGAGTTCAGACATCCCTTTTTGTATGCGGGAGTCTGATCTCTGCGGATAGGCT	2340
Db	2281	AAACCAAGAGTTCAGACATCCCTTTTTGTATGCGGGAGTCTGATCTCTGCGGATAGGCT	2340
Qy	2341	AAACCAATTAATATGAAAAATTTCTATTTAAAGTCACAGAAAAATTTATGCGTGTATATCAAA	2400
Db	2341	AAACCAATTAATATGAAAAATTTCTATTTAAAGTCACAGAAAAATTTATGCGTGTATATCAAA	2400
Qy	2401	TTTGGGGAAATTTCTTGTAACCAAAAGGAAAAATTAATCTTGCTTTGGCTTCACAGAA	2460
Db	2401	TTTGGGGAAATTTCTTGTAACCAAAAGGAAAAATTAATCTTGCTTTGGCTTCACAGAA	2460
Qy	2461	ACTCATTTGCTTGAAGTCGAGAAAGTATCTCTCAAAAATCTTAAGTCTCTAAATTAAC	2520
Db	2461	ACTCATTTGCTTGAAGTCGAGAAAGTATCTCTCAAAAATCTTAAGTCTCTAAATTAAC	2520
Qy	2521	AGAGCTGAAACTTAAAGGCAAGCTGCAGATTAAGTTGATGCTATGAGATTTTGAACCTT	2580
Db	2521	AGAGCTGAAACTTAAAGGCAAGCTGCAGATTAAGTTGATGCTATGAGATTTTGAACCTT	2580
Qy	2581	TAGTATATTAGTTCATGATTAATTAAGCAATGCCATAGATTAATCCCTCAACAGCAATAAATTA	2640
Db	2581	TAGTATATTAGTTCATGATTAATTAAGCAATGCCATAGATTAATCCCTCAACAGCAATAAATTA	2640
Qy	2641	AGTGCACATGAAAAAAAAGCCAGACTTAAACAAGAAAAAAGTTGCAAAAACATCATCAAA	2700
Db	2641	AGTGCACATGAAAAAAAAGCCAGACTTAAACAAGAAAAAAGTTGCAAAAACATCATCAAA	2700

QY 2701 AGAGATTGAGTTAACCTGATGTTAAAGACACATTTTGTGTAAGAAAGATGTA 2760
DB 2701 AGAGATTGAGTTAACCTGATGTTAAAGACATTTTGTGTAAGAAAGATGTA 2760
QY 2761 TTTTCAAGAGTGATACATTTATGCTTTTTCAGGGATCTTTCAAGAAAAGTCCTTTG 2820
DB 2761 TTTTCAAGAGTGATACATTTATGCTTTTTCAGGGATCTTTCAAGAAAAGTCCTTTG 2820
QY 2821 GGGGTACAGGAAGCTTAGAAAACATTTGAAAGATGAAAATGAGCAAAATTAAGAAAAT 2880
DB 2821 GGGGTACAGGAAGCTTAGAAAACATTTGAAAGATGAAAATGAGCAAAATTAAGAAAAT 2880
QY 2881 GGTTTTACCAAGGACATGGAATCTTTACTTTGCAATTAATTTTATTTGCTCTTTT 2940
DB 2881 GGTTTTACCAAGGACATGGAATCTTTACTTTGCAATTAATTTTATTTGCTCTTTT 2940
QY 2941 CTCTAGCTAACAGACCTTAAAGAAAGCATGTTCATGACTTCAATTTCACTCCAGGGGA 3000
DB 2941 CTCTAGCTAACAGACCTTAAAGAAAGCATGTTCATGACTTCAATTTCACTCCAGGGGA 3000
QY 3001 CTTCAGAGCTGCCGTCGCCAGAAAGTTCAAGAGTAAAGGCGACGAGTAACTGATGTT 3060
DB 3001 CTTCAGAGCTGCCGTCGCCAGAAAGTTCAAGAGTAAAGGCGACGAGTAACTGATGTT 3060
QY 3061 TTTTCAAGAAAGAGATCTGAGGCGAGGTTTCACTGGGTGATGAGACAGCTGATGTT 3120
DB 3061 TTTTCAAGAAAGAGATCTGAGGCGAGGTTTCACTGGGTGATGAGACAGCTGATGTT 3120
QY 3121 AAGAAATTTTATGTGAGCTTGTCTACGAGGCGGCTTCTCTTCAAGGCAATTTACT 3180
DB 3121 AAGAAATTTTATGTGAGCTTGTCTACGAGGCGGCTTCTCTTCAAGGCAATTTACT 3180
QY 3181 GAGCGAGCTTCCCAAGTCTGCTCTGCGCAATGCTGTCTAATTTCCCTGGGAAAAAAGT 3240
DB 3181 GAGCGAGCTTCCCAAGTCTGCTCTGCGCAATGCTGTCTAATTTCCCTGGGAAAAAAGT 3240
QY 3241 CAACATTAATAAAAGTCTCTCTCTCCCTTCAAGGCTCTTCCCTTCCCACTCC 3300
DB 3241 CAACATTAATAAAAGTCTCTCTCTCCCTTCAAGGCTCTTCCCTTCCCACTCC 3300
QY 3301 CCTAGAGCAGAGAGAGAGCTCCCTTTTGGGCGAGCTCATCTTACTTGAATTTGAGAA 3360
DB 3301 CCTAGAGCAGAGAGAGAGCTCCCTTTTGGGCGAGCTCATCTTACTTGAATTTGAGAA 3360
QY 3361 GCTGGGTGAAGCTCTTATGCGACAGTTTCAAGGGATTTAGCAGGTGAGTGAACATAG 3420
DB 3361 GCTGGGTGAAGCTCTTATGCGACAGTTTCAAGGGATTTAGCAGGTGAGTGAACATAG 3420
QY 3421 CTGGGAGAGACTTGAATGAGAGTCCGCCCCCAATTTCAATTTTAAAGCAGGT 3480
DB 3421 CTGGGAGAGACTTGAATGAGAGTCCGCCCCCAATTTCAATTTTAAAGCAGGT 3480
QY 3481 GAGACATCATAGAAAGTTCATAGCACTCAGGACCTGTGCAAGACACATGCGACAGGA 3540
DB 3481 GAGACATCATAGAAAGTTCATAGCACTCAGGACCTGTGCAAGACACATGCGACAGGA 3540
QY 3541 GAGAGACATGATTAACCTTAAACAGCTTGAAGAAAACAACTGCTGCTTAATTTAA 3600
DB 3541 GAGAGACATGATTAACCTTAAACAGCTTGAAGAAAACAACTGCTGCTTAATTTAA 3600
QY 3601 AATCAGCCCACTTAAATGTTTATCAGCTTCCCTTCTGATTTCAATTCAGGAATTC 3660
DB 3601 AATCAGCCCACTTAAATGTTTATCAGCTTCCCTTCTGATTTCAATTCAGGAATTC 3660
QY 3661 AAGAAATAGACATTTCTACTACTGACCCAAAGAACAATTTACTCTTCAAGGCTGTG 3720
DB 3661 AAGAAATAGACATTTCTACTACTGACCCAAAGAACAATTTACTCTTCAAGGCTGTG 3720
QY 3721 GAGGCGACAGTTGGTAAAGGCTCTTACAGGTTTTTATATCCCTCAATCACAAT 3780
DB 3721 GAGGCGACAGTTGGTAAAGGCTCTTACAGGTTTTTATATCCCTCAATCACAAT 3780

QY 3781 GACAGATTTTGTATATGCGAAGCTGGAATTTGCTGCTTCAATTCCTCACCTGCTTAT 3840
DB 3781 GACAGATTTTGTATATGCGAAGCTGGAATTTGCTGCTTCAATTCCTCACCTGCTTAT 3840
QY 3841 AAGAAACTGAAGTTGGTTTTCTGCAAAATTAATGTACATGCAAAAGATGATTAATCTAG 3900
DB 3841 AAGAAACTGAAGTTGGTTTTCTGCAAAATTAATGTACATGCAAAAGATGATTAATCTAG 3900
QY 3901 ATTTTATATTTGCAAAATACACAAATGCTGGAATTAATAATCTGTTATTCAAA 3960
DB 3901 ATTTTATATTTGCAAAATACACAAATGCTGGAATTAATAATCTGTTATTCAAA 3960
QY 3961 AGCTAAGTACTAATTTGGTAAACAACCACTTGTAAATATATGTAAAGATCAGTA 4020
DB 3961 AGCTAAGTACTAATTTGGTAAACAACCACTTGTAAATATATGTAAAGATCAGTA 4020
QY 4021 ATTTCCCTTTTATGTCAGAGTGGGAAAGTGAATGCTGCTTTTCTTTATGTTACTCA 4080
DB 4021 ATTTCCCTTTTATGTCAGAGTGGGAAAGTGAATGCTGCTTTTCTTTATGTTACTCA 4080
QY 4081 ATAGAGAAAGAAATATGCTCAATATGCTTAATATTTAATTTAAATAATAGCTGAT 4140
DB 4081 ATAGAGAAAGAAATATGCTCAATATGCTTAATATTTAATTTAATAATAGCTGAT 4140
QY 4141 CCGAGTGCAGTGTGTTTACACTACTTGATCACAACAGTTACAGATTTCTTGTCTCT 4200
DB 4141 CCGAGTGCAGTGTGTTTACACTACTTGATCACAACAGTTACAGATTTCTTGTCTCT 4200
QY 4201 TCTCCACTCCCACTGCTTCACTTAACTGCGCAAAAAGAAAAAGAAAAATTTATATA 4260
DB 4201 TCTCCACTCCCACTGCTTCACTTAACTGCGCAAAAAGAAAAAGAAAAATTTATATA 4260
QY 4261 CTACTCAAGACTAAATATTTATTTATTTACTAGTATTTATAGCTTATTTATTTT 4320
DB 4261 CTACTCAAGACTAAATATTTATTTATTTACTAGTATTTATAGCTTATTTATTTT 4320
QY 4321 ACTTGTTAAACAGGATTTGAGGAGACATACAGTTTATTTATTTATTTATATAT 4380
DB 4321 ACTTGTTAAACAGGATTTGAGGAGACATACAGTTTATTTATTTATTTATATAT 4380
QY 4381 TTAATTTATTTTGAATGGAATCTGTGACCCACGCTGAGTGCAGTGTGATC 4440
DB 4381 TTAATTTATTTTGAATGGAATCTGTGACCCACGCTGAGTGTGATGATC 4440
QY 4441 TCAATGATGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500
DB 4441 TCAATGATGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4500
QY 4501 TACTTTCTTAATCTGTTTATGATGAGTGTACCTTAATCTTTTCAATGCTGAGAACATC 4560
DB 4501 TACTTTCTTAATCTGTTTATGATGAGTGTACCTTAATCTTTTCAATGCTGAGAACATC 4560
QY 4561 TGCATTAAGAGACACATTTTATTTTATCTTGAAGCTTCTCATATCATTTGGCCATG 4620
DB 4561 TGCATTAAGAGACACATTTTATTTTATCTTGAAGCTTCTCATATCATTTGGCCATG 4620
QY 4621 AACTGTTTTCAAGTGGCTGGAACGCGGGGACCCCTGGAACATCTTGGATCATGGGCA 4680
DB 4621 AACTGTTTTCAAGTGGCTGGAACGCGGGGACCCCTGGAACATCTTGGATCATGGGCA 4680
QY 4681 CCATGACACTTCTGATCTCTCTCTGAGTTCGACTTGAATGTTGTGCAAGACCTT 4740
DB 4681 CCATGACACTTCTGATCTCTCTCTGAGTTCGACTTGAATGTTGTGCAAGACCTT 4740
QY 4741 TCCAGCCCGAAGTTTACACAGAAATTCATTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4800
DB 4741 TCCAGCCCGAAGTTTACACAGAAATTCATTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4800
QY 4801 TTTTCAATTTAATCTTCTTCACTACAGGAAATTTATATGATTTCAACAATCACCAC 4860
DB 4801 TTTTCAATTTAATCTTCTTCACTACAGGAAATTTATATGATTTCAACAATCACCAC 4860
QY 4861 TCCATTTATTAATTTCTTTCTGATGTTGAAATGCTGCATGATTAATATATAG 4920

DB 4861 TCCATTATTAAGATTTCTTCTCGATGCGTTGAAAGCTCCAGATTATATATAG 4920
QY 4921 ATCTCAGAACTGTAATCTTCTGTTCTATCTTTTAAATCATGTTCTTAAT 4980
DB 4921 ATCTCAGAACTGTAATCTTCTGTTCTATCTTTTAAATCATGTTCTTAAT 4980
QY 4981 CTAATCTTTCTTAATTTGT 5000
DB 4981 CTAATCTTTCTTAATTTGT 5000
RESULT 2
AAK71358/c
ID AAK71358 standard; DNA; 37314 BP.
XX
AC AAK71358;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26170.
XX
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.

PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PA
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 XX
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis
 XS
 XS Disclosure, SEQ ID NO 26170; 3071pp + Sequence Listing, English.

CC amino acid sequences given in AAM82170 to AAM91921. (1) have cytosolic
CC AAK5495 to AAK64702 encode the human immune/haematopoietic antigen (1)
CC activity and can be used in gene therapy and vaccine production. (1)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (1) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (1) by expressing inactive proteins or to
CC supplement the patients own production of (1). Additionally (1)
CC polynucleotides may be used to produce the secreted (1), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (1) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK54703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM22169
CC represent sequences used in the exemplification of the present invention
CC
CC
CC Sequence 37314 BP; 8700 A; 8411 C 9616 G; 10587 T; 0 U; 0 Other;
SQ

	Query Match	Best Local Similarity	7.4%; Score 370.6; DB 4	Length 37314;
	Matches 639;	Conservative 0;	Mismatches 249;	Indels 49; Gaps 7;
QY	11	TGCTGTGGCTCAGCGCGTATATCCACACTTTGGGAGGC-----AGAGAGATCGCTTG	64	
Db	28221	TGCGGTGGCTATGCCCTGTATATCCACACTTTGGGAGGCCAAGTGGGTGCACCTGG	28162	
QY	65	AGCTCAGGAATTGGAGACAAGCTTACGTAACATAGTAAACCTGTCTGTATACAAATTAAT	124	
Db	28161	AGGTAGGAGATTGAAACCAAGCTGGCCCAACATGGTAAAC-C-CTGCTCTTATTAATAAT	28103	
QY	125	AAAGAAATTTTCCAGGATGCTGGCCGCAACCCCGATGTCACCTATTGGGAGGCTGAG	184	
Db	28102	ACAAAATTTAGCCAGGCGTGTGCTCAACCTGTATATCCAGCACTATGGAGGCCGAG	2804	
QY	185	GTAGAGGAATGCTTGAAGCCAGAGTTGAAGACAGCTTAGGCAACATAGTAGACCTT	244	

QY	245	GTGCTATAATAA-----AATTAATAGCTGGTGTGCTTGGACACAGGCGCTGCA--GCTAGCTA	298
Db	28042	ATGGGTGGAATCTCTTGAGGTCAGGAAGTTATGATGCCAGCTGGCCAACTATGTGAACCCC	27983
QY	27982	ATCTCTACTATAAATACAAAAATTAGCCGGGCGCTGTGTGGGCGATGCTGTAAATCCACGCTA	279223
QY	299	CTCGAAGAAGCTGAGGTGGAGGA-TCATGAGCCCGAGAGGCTGAGGCTGCAGTGAACAG	357
Db	27922	CTCAGAGGCGCTGAGCAGAGGAATGTGTGAACCCCGAGAGGTGAGGTTGCAGTGAAGCCA	27863
QY	358	TGATTCACCCAGCTGGAATT-----CGAGCGTGAAGAACAAGGGAGAGCCCGTTTCCAAAA	412
Db	27862	AGATCACACCATGCACTGCACTGCACCTGGCGCAGGAGCAAGCTTCATCTCAAAAA	27803
QY	413	AAAAAAAAAAAAAAAAATGCAGAAAAAGACATCATATACTTGAACCTGGAGACATACTTTT	472
Db	27802	AAAAAAAAAACACACAAAAAAACAAATTAAGCAGAGSCATGTGGCTACTCAGACAGCTGA	27743
QY	473	ATGTGATGAATTCACAAATCTTTTAGAAGAAATTAGCAATTTCTGTATAAATGTATTA	532
Db	27742	GGCAGAGGAATCACTTGAACCCAGAGGGCGAGGTGCAGTAGACCAATCGTGCACT	27683
QY	533	ATTATATTATTATAAATTCAAATGGAATTAAATATTCTGAGAAACTAGCTTCTCACTTC	592
Db	27682	GCACTCCAGCTGGGTGACAGAGTAAGACTCCGCTCTCAAAAA-----	27640
QY	593	TCAGTTGTACGTCAAAACCTTAATGCTCTTTGGCCGGGTGCGGTGCTCACGCTGTAA	652
Db	27639	-----AAAAAAAAAAAAAAAAATTTTGGCCAGGCGGTGGCTCACGCTGTAA	27593
QY	653	CCGAGCACTTTGGAGAGCCGAGGCGGGTGGATCACAAAGTTTGGAGATGAGACCATCT	712
Db	27592	CCCCGCACTTTGGAGAGCCAAACGGGCGGATCACAGGTCAAGAGCTGAAAGCAATCTT	27533
QY	713	GGCTAAACACGCGTAAACCTCGTCTCTACTATAAATACAAAAAATTGACCGGGTGCCTGC	772
Db	27532	AGCTAACACGGTGAAACCCCGTCTTACTATAAAATACAAAAATTGCGGGCATGTGG	27473
QY	773	CAGACGCTGTATGCCAGCTGCTCAGAGAGGCTGAGGACAGAGAAATGTTGTAAACCCGGG	832
Db	27472	CAGCGCGCTGTATGCCAGCTACTCAGAGGCTGAGGACAGAGAAATGTTGTAAACCCGGG	27413
QY	833	AGGCGAGACTTGGAGTGAAGCCGAGATTGCGCCTCACTTCACAGCTTGGGCGACAGTGGG	892
Db	27412	AGGCGAGATTGAGAGTGAAGCCGAGATGGCGCCACTGCACTCAAGCCTGGGCGACAAAGG	27353
QY	893	AGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAGTTGA	929
Db	27352	AGACTCGTCTCAAAAAAAAAAAAAAAAAAAATTA	27316
RESULT 3			
ABQ88207 standard; cDNA, 172570 BP.			
ID ABQ88207			
ABQ88207;			
XX	AC		
XX	DT	18-SEP-2002 (first entry)	
XX	DE		
XX	XX	Human osteoblast differentiation related cDNA SEQ ID NO 114.	
XX	KW	Human; osteoblast; stem cell differentiation; bone tissue deposition;	
XX	XX	osteoporosis; osteopathic; ss.	
XX	OS	Homo sapiens.	
XX	XX		
XX	PN	MO200250301-A2.	
XX	PD	27-JUN-2002.	
XX	PF	18-DEC-2001; 2001MO-US048276.	
XX	PR	18-DEC-2000; 2000US-0255882P.	

PR 24-APR-2001; 2001US-0286591P.
 XX (GENE-) GENE LOGIC INC.
 PA (PROC) PROCTER & GAMBLE CO.
 XX Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A,
 PI Metz L,
 XX WPI; 2002-557663/59.
 DR
 XX
 XX Use of genes and their expression profiles associated with osteoblast
 PT differentiation for screening modulators bone formation, for diagnosing
 PT or treating e.g. osteoporosis, or as markers for the differentiation
 PT process.
 XX
 XX Claim 1; SEQ ID NO 114; 78pp + Sequence Listing; English.
 XX
 XX The invention relates to genes and their expression profiles are used
 CC for: (a) screening modulators of precursor stem cell differentiation into
 CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
 CC deposition of bone tissue, abnormal rate of osteoblast formation or
 CC osteoporosis; or (c) treating or monitoring treatment of the conditions
 CC cited in (b), or monitoring the progression of bone tissue deposition.
 CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
 CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-
 CC induced abnormalities in bone formation or bone loss, conditions that
 CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
 CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
 CC or fibrous dysplasia. The present sequence is that of an osteoblast
 CC differentiation associated cDNA marker of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_poc_sequences
 XX
 XX Sequence 172570 BP; 45194 A; 43991 C; 41126 G; 42258 T; 0 U; 1 Other;
 SQ

Query Match 7 4%; Score 370.6; DB 6; Length 172570;
 Best Local Similarity 68.2%; Pred. No. 1.9e-51;
 Matches 639; Conservative 0; Mismatches 249; Indels 49; Gaps 7;

QY 11 TCGTGTGCTCAAGCGGTATTCCTCCAGCACTTTGGAGAGC-----AGGAGATCGCTTG 64
 DB 136255 TCGCGTGTCTATGCTCTGTAATCCAGCACTTGGAGAGCGCAAGGTGGATCACTCG 136314
 QY 65 AGCTCAGAAATTGAGACAAAGCCTAGTAACATAGTGAACCTCTGCTGTAACAAATAAT 124
 DB 136315 AGGTGAGAGTGAACAAACAGCCTGGCCAAACATGTGAACCC-CTGTCTCTACTAAAGT 136373
 QY 125 AAAAAGATTTTCCAGGATGTGTGGCTGACCCCGCAGCTATTTGGAGAGCTGAG 184
 DB 136374 AAAAAAATTAGCCAGGCGTGGGCTCAACACCTGTATCCAGCACTTATGGAGGCGCGAG 136433
 QY 185 GTAGAGAGAAATGCTTGAAGCAGAGAGTTGAAGCAAGCCTAAGCAATATGAGACCT 244
 DB 136434 ATGGGTGATCTCTTGAAGTCAAGAGTTATACACAGCTCGCAACATGTGAACCC 136493
 QY 245 GTGTCTATTAATA-----AATTAATTAAGTGTGTCTTGGACAGAGCTGCA--GCTAGCTA 298
 DB 136494 ATCTCTACTAAATAATCAAAATAATTAAGCCGGGCGGTGGGCAAGCTGTAAATCCAGCTA 136553
 QY 299 CTCGGAAGACTGAGGTGGAGGA--TCACTGAGCCCGAGAGGCTGAGGCTGCACTGAACAG 357
 DB 136554 CTCAGAGGCTAGAGGAGAGAAATGTTGAACCCCGAGAGGTGAGAGTTCAGGTGAGCA 136613
 QY 358 TGATCAACCCAGCTGAT-----CCAGCCTGGAAGACAGAGGAGACCTGTTCCAAA 412
 DB 136614 AATTCACACATCTGCACTGCAATCACTGCGGCGAGCAAGCACTCATTTCAAAA 136673
 QY 413 AAAAAAAAAAAAAAAAAATGCAAGAAAAGACATCATTAACCTTGACTGGAGCATTAACCTTTT 472
 DB 136674 AAAAAAAAAAACAACAAAAAACAATTAGCAGAGCAATGTGTGCTACAGAGACTGA 136733
 QY 473 ATGTGATGAATTCAATCTTTTAGAGAAAGAAATTAGCATTTCTGTATAAATGATATA 532

DB 136734 GGCAGAGAAATCACTTGAACCCAGAGAGCGGAGGTTGCACTAGCCAAATGTGCACT 136793
 QY 533 ATTATATTATTATTAATTAATTCAAATGAATTAATTAATTTCTGAGAAACTAGCTTCACTTC 592
 DB 136794 GCATTCACAGCTGGGTGACAGAGTAAGTCCGCTCAAAAA----- 136836
 QY 593 TCAAGTTTCAGTCAAAACCTTTAATGTCTTTGCGCGGCGCGGTGCTCAAGCCTGTAT 652
 DB 136837 -----AAAAAAAAAAAAAAAAATTTTGGCAAGGCTGTGCTCAAGCCTGTAT 136883
 QY 653 CCAGACCTTTGGAGAGCGCGAGGCTGATCAACAGTTAGAGATGAGACCACTCT 712
 DB 136884 CCCCACCTTTGGAGAGCGCAAGCGGCGGATCAAGAGTCAAGAGTGAACATCTT 136943
 QY 713 GGCTAACAGAGTGAACCTGCTCTACTTAATAAATCAAAAAATTAAGCCGGGTGCGTGC 772
 DB 136944 AGCTAACAGGTGAACCCCGTCTCTACTTAATAAATTAAGCCGGGCATGTGC 137003
 QY 773 CAGAGCGCTGTAGTCCCACTGCTCAAGAGGCTGAGGAGAAATGCTGTGAACCCGGG 832
 DB 137004 CAGGCGCTGTAGTCCCACTGCTCAAGAGGCTGAGGAGAAATGCTGTGAACCCGGG 137063
 QY 833 AGCGAGGCTTGCAGTGAAGCCGAGATTGGCGCCACTGCACTCCAGCCTGGGCGACAGTGC 892
 DB 137064 AGCGGAGTTTCACTGAGAGCCAGATGGCGCACTGCACTCAAGCCTGGGCGACAAAGCG 137123
 QY 893 AGACTCTGTCTCAAAAAAAGTTGA 929
 DB 137124 AGACTCCGTCTCAAAAAAAGTTGA 137160

RESULT 4
 AAK76548/C
 ID AAK76548 standard; DNA; 13673 BP.
 XX
 AC AAK76548;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:31360.
 XX
 KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 XX cytostatic; gene therapy; vaccine; metacastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001354.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-022824P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235634P.
PR 27-SEP-2000; 2000US-0235636P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241211P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0250306P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0251989P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 31360; 3071bp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX

Seq Sequence 13673 BP: 3494 A; 2746 C; 3372 G; 4061 T; 0 U; 0 Other;

Query Match 7.0%; Score 351.8; DB 4; Length 13673;

Best Local Similarity 67.1%; Pred. No. 2e-48;

Matches 626; Conservative 0; Mismatches 282; Indels 25; Gaps 8;

```
13 CTGTGCTCAGCGGTGATTTCCAGACATTTGGGAGCG-----AGGAGATCGCTGAG 66
DB CAGTGGCTCAGCTGTATATCCAGACATTTGGGAGCGCAAGGTGGTGTATCTACTAG 7827
QY 67 CTGAGGAATGGAGCAAGACCTACGTATAGTGAACCTCTGTCTGTACAAATATATA 126
DB 7826 GTGAGAGTTAGAGCAAGCTCTGGCCAAAGGGTGAACCCCATCTCTATAAATAGA 7767
QY 127 AAGAAATTTCCAGGATGTGTGGCGTGCACCCCAAGTGCAGCTATTTGGGAGCGTGA 186
DB 7766 AAAAAATGAGCCAGGATGTGTGGCTGCACCTGTATATCCAGACATTTGGGAGCGCAG 7707
QY 187 AGGAGGAATGCTTGAAGCCAGATTTGAAGACAGCAAGCTTAGTGAACCTCTGT 246
DB 7706 AGGTG--AATCAAGAGTCAGAGATCAAGACCATCTGCTTAACAGGTGAACCCCT 7649
QY 247 GCTATATA-----AAAAATAATAGCTGTGTCTTGGCAGAGCGCTGAG--CTAGCT 297
DB 7648 TCTACTAAAAATACAAAAAAATTTAGCCGGCATGTGTACAGCGCTGTATCCAGCT 7589
QY 288 ACTCGAAGACTGAGGTGGAGAGATCAC--TGAGCCAGAGGCGTGAAGCTGAGTGAACA 356
DB 7588 ACTTAGAGAGCTTAAGCAGAGAGATCACTTGAACCTGAAGAGAGAAAGCTGAGGGAGCG 7529
QY 357 GTGATACCCAGCTGATTTCCAGCTGGAAGACAGAGGAGACCTGTTTCCAAAAA 416
DB 7528 GAGATACCGCCACTACCTCCAGCTGGCGGAGCAGAGGAGAGCTGTGTACAAAAA 7469
QY 417 AAAAAAATAATCAAGAAAGACATCATTAATCTTGAAGCTGAGCATTAATCTTATGT 476
DB 7468 AAAAAAACAAGAGCTGCTATTAGATTTGAACCAAGTTCACTCATATATAAGTAC 7409
QY 477 GATGAATTCACAACTTTTAGAAGAAATTAAGCATTTCTGATTAATTAATTAATTA 536
DB 7408 AAAAAAGTTACTTCAAGTATAGTCAAAAGTCAAAATCTTCAATTAATAGACAGGCGCACA 7349
QY 537 TATTTATTAATTAATTAAGAAATTAATTAATCTGAGAACTAGCTTCTCA-----CTCTC 592
DB 7348 AAAAAAACAAGAGGATCATTTTGAAGCAAAATGATGAGCACTTCAACTGTCTAT 7289
QY 593 TCAGTTGTCAGTCAAAACTTTAATGATCTTTGGCGGGTGGCGGTGCTCAGCGCTGTAT 652
DB 7288 TCAATTTCTCTTAAGAAATTTTCTTCTGTTGGTGGTGGCTGAGCGCTGTAT 7229
QY 653 CCAGCACTTTGGAGCGCGGAGGCGGTGATCACAAAGTTAGAGATGAGACCAATCT 712
DB 7228 CCAGCACTTTGGAGCGCGGAGGCGGTGATCATGAATGAGAGATCAAGCAATCTCT 7169
QY 713 GGCTTAACAGGTGAACCTCTCTCTACTTAATAAATAAATAAATAAATTAAGCGGGTGG 772
DB 7168 GGCTAATAGGGTGAACCAATCTCTACTTAATAAATAAATAAATAAATAAATAAATAA 7110
QY 773 CAGAGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGAGGAGAAATGATGTGA--CCCG 830
DB 7109 TGGGGCGCTGTAGTCCAGCTACTCGGAGGCTGAGGAGGAAATGATGAAGCTTCG 7050
QY 831 GGAGCGGAGCTTGAAGTGAAGCGAGATTCGCACTGCACTCCAGCTGAGCGAGTGA 890
DB 7049 GGAGGTGAGCTTGAAGGAGCTGAGATGTGCACTGCACTCCAGCTGAGCGTGAAGAG 6990
QY 891 CGAGACTCTGTCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 923
DB 6989 CAAAGACTTGTCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6957
```

RESULT 5
AAK81195/c

```
ID AAK81195 standard; DNA; 13673 BP.
XX
XX AAK81195;
AC AAK81195;
XX
XX 07-NOV-2001 (first entry)
DT
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36007.
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
XX NC_000015:17182-17182.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216800P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 11-JUL-2000; 2000US-0217496P.
XX
XX 14-JUL-2000; 2000US-0218290P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
XX
XX 14-AUG-2000; 2000US-0225267P.
XX
XX 14-AUG-2000; 2000US-0225268P.
XX
XX 14-AUG-2000; 2000US-0225270P.
XX
XX 14-AUG-2000; 2000US-0225477P.
XX
XX 14-AUG-2000; 2000US-0225758P.
XX
XX 14-AUG-2000; 2000US-0225759P.
XX
XX 18-AUG-2000; 2000US-0226279P.
XX
XX 22-AUG-2000; 2000US-0226881P.
XX
XX 22-AUG-2000; 2000US-0226882P.
XX
XX 22-AUG-2000; 2000US-0227182P.
XX
XX 23-AUG-2000; 2000US-0227183P.
XX
XX 30-AUG-2000; 2000US-0228924P.
XX
XX 01-SEP-2000; 2000US-0229287P.
XX
XX 01-SEP-2000; 2000US-0229343P.
XX
XX 01-SEP-2000; 2000US-0229344P.
XX
XX 01-SEP-2000; 2000US-0229345P.
XX
XX 05-SEP-2000; 2000US-0229509P.
XX
XX 05-SEP-2000; 2000US-0229513P.
XX
XX 06-SEP-2000; 2000US-0230437P.
XX
XX 06-SEP-2000; 2000US-0230438P.
XX
XX 06-SEP-2000; 2000US-0231242P.
XX
XX 08-SEP-2000; 2000US-0231243P.
XX
XX 08-SEP-2000; 2000US-0231244P.
XX
XX 08-SEP-2000; 2000US-0231413P.
XX
XX 08-SEP-2000; 2000US-0231414P.
XX
XX 08-SEP-2000; 2000US-0232080P.
XX
XX 08-SEP-2000; 2000US-0232081P.
XX
XX 12-SEP-2000; 2000US-0231968P.
XX
XX 14-SEP-2000; 2000US-0232397P.
XX
XX 14-SEP-2000; 2000US-0232398P.
```

[illegible]

Query Match	Best Local Similarity	Score	DB	Length
Matches 626; Conservative 0; Mismatches 282; Indels 25; Gaps 8;	7.0%;	351.8;	4;	13673;
13 CTGTCGCTCAGCCGTGATTTCCCGACCTTTGGGAGGC-----AGGAGGATGCGCTTGAG	66			
DB 7886 CAGTGGCTCCACACTTGATATCCCGACCTTTGGGAGGCCTGAGGTGGTGAATCACTTGAG	7827			
QY 67 CTGAGGAATTTGAGACCAAGCCTTACGTAACTATGTGAACCTCTGTCTGTACAAATATATAA	126			
DB 7826 GTGAGGAATTTGAGACCAAGCCTTACGTAACTATGTGAACCTCTGTCTGTACAAATATATAA	7767			
QY 127 AAGAAATTTTCAGGAGATGTCGTCGACCCCGACGTGACAGCTATTTTGGAGGCTGAGGT	186			
DB 7766 AAAATTTAGCCAGGAGATGTCGTCGACCCCGACGTGACAGCTATTTTGGAGGCTGAGGC	7707			
QY 187 AGAGGAATGCTTTGAAGCCAGAGATTTGAAGCAAGCCTTACGATATAGTGAAGCCCTGT	246			
DB 7706 AGGTG--AATCAAGAGGTTCAGAGATCAAGCACTCTGCTTACACGGTGAACCCCGT	7649			
QY 247 GCTATATA-----AAAATATATTAGCTGTGTTCTTTGGCAAGGCTGACG--CTAGCT	297			
DB 7648 TCTACTTAAAAATACAAAAAAATTTAGCCGGGCAATGGTGAACACGGGCTGTATGTCAGCT	7589			
QY 298 ACTCGGAAGACTGAGTGGAGAGATCAC--TGAGCCAGAGAGCTGAGGCTTGCACTGAACA	356			
DB 7588 ACTTAGAGAGGCTTAAAGCAGAGAACTTGAACCTGAGAGGCAAGACTTGCAAGGAGCC	7529			
QY 357 GTGATACACCACTGAGATTTCCAGCTGGAACAACAGAGGAGAACCTGTTTCAAAAAAAA	416			
DB 7528 GAGATACAGCCCACTTACCTCCAGCTTGGGCAACAGAGGAGAACTGTGTACCAAAAAAAA	7465			


```
QY 417 AAAAAAAAAAATGCAAGAAAGACATCAATACTGAGGACATACTTTATGT 476
DB 7468 AAAAAAAAAAAGAGACTGCTATTATGTTGAAACCACTTCATCATTAAAGTAC 7409
QY 477 GATGAATTCACATCTTTTGAAGAAATAGCATTTCTGATTAATGATTATTA 536
DB 7408 AAAACGGTACTCAGCTAGTCAGCAAGCAATCTTCAATTAATGCTGAGGCAACA 7349
QY 537 TATTATTATTAATTCATTAATTAATTTCTGAGAAATGACTTCA---CTTC 592
DB 7348 AAGGAAACAAAGGGGATCATATTGAGCAAAATGATGCAAGCACTTCACTTGTCAT 7289
QY 593 TCAGTTCAGTCATCAAACTTTAATGCTTTGGCCGGGTCGGATGCTACGCTGTAAT 652
DB 7288 TCATTTCTCTTAAGAAATTTTCTCTTCTGTTGGGTCGGGCTCAAGCCTGTAAT 7229
QY 653 CCCAGCATTTGGAGAGCCGAGGCGGGTGATCAAAAGTTAGAGATCGAGACATCT 712
DB 7228 CCCAGCATTTGGAGAGCCGAGGCGGGTGATCATGAAGTCAGAGATCAAGACATCT 7169
QY 713 GGTAAACAGGTGAAACCTGCTCTAATAAATAAATAAATTAAGCCGGTGGCTGC 772
DB 7168 GGTAAATACGTGCAACACCATCTCTAATAAATAC-AAAAAATAATTAAGCAGGTG 7110
QY 773 CAGAGCGCTGTAGTCCAGCTGCTCAAGAGCTGAGGAGAGAAATGCTGTA--CCCG 830
DB 7109 TGGGCGCTGTAGTCCAGCTACTCGGAGGCTGAGGAGAGAAATGAGCAAGCTCCG 7050
QY 831 GGAGGCGAGCTTGCAGTGAAGCCAGATTGGCCCACTGCATCAGCCTGGGAGAGT 890
DB 7049 GAGAGGAGAGCTTGCAGGCGAGCTGAGATCGTGCATCGCATCAGCCTGGGAGAG 6990
QY 891 CGAGACTCTGTCTCAAAAAAAAAAAAAAAAAA 923
DB 6989 CAGACTTGTCTCAAAAAAAAAAAAAAAAAA 6957

RESULT 6
ADL13909/c
ID ADL13909 standard; DNA; 129588 BP.
AC ADL13909;
XX
DT 06-MAY-2004 (first entry)
DE Osteoarthritis-associated polymorphic nucleotide #441.
XX
KM de; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KM joint space narrowing; osteophyte development; joint pain;
KM osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
PN WO2003054166-A2.
XX
PD 03-JUL-2003.
PF 19-DEC-2002; 2002WO-US041225.
PR 20-DEC-2001; 2001US-0342603P.
XX
PA (INCY-) INCYTE GENOMICS INC.
PI Jones KA, Schaffer A;
DR WPI; 2003-559141/52.
XX
PT Determining susceptibility of an individual to joint space narrowing,
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PS a protein.
PS Disclosure; SEQ ID NO 441; 297bp; English.
```

```
XX The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space
CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).
SQ Sequence 129588 BP; 35710 A; 27530 C; 26424 G; 36620 T; 0 U; 3304 Other;

Query Match 7.0%; Score 351; DB 10; Length 129588;
Best Local Similarity 66.6%; Pred. No. 3e-48;
Matches 623; Conservative 0; Mismatches 275; Indels 37; Gaps 7;

QY 1 TATAGGCCAATGCTGTGCTCAAGCGCTGATTTCCAGCACTTGGAGGC---AGAGGA 57
DB 23050 TAAAGCCTGTGTCAGTGTGCTCACTCAATCCAGTATTTAGAGGCAAAAGCAGA 22991
QY 58 TCGCTTGAGCTCAGGATTTGAGACAAAGCCTAGTAACTAGTGAACCTGTCTGTAC 117
DB 22990 TCACCTTGAGCTCAGGATTTGAGATGAGCTGGGCAACATAAGAGATCCATCTTACA 22931
QY 118 AATAATTAAGAATTTCCAGGCAATGTCGTCGTCGACCCCAAGTCAGCTAATTTGGGA 177
DB 22930 AAAAATAATTAAGTGTGGGCAAGTGTGCTCACTGCTAATATCTCACTTTGGGA 22871
QY 178 GGCTGAGTGAAGAGAACTCTGAAGCCGAGGTTGAAGACAGCCTAGCAACATAGTG 237
DB 22870 GGCTGAGTGAAGAGAACTCTGAAGCCGAGGTTGAAGACAGCCTAGCAACATAGTG 22811
QY 238 AGACCTGTGTCTAATAAATAAAT---AATAGCTGTGTGTCTTGGCAGAGCCTGCA-- 290
DB 22810 AATCTCATCTCTACTAATAAATAAATAAATAAATTAACAGGATGTGGCAGACCTGTAT 22751
QY 291 GCTAGCTACTCGAAGACTGAGGTGGAGATCAC-TGAGCCAGAGAGCTGAGGCTGCA 349
DB 22750 CCAGCTACTTGGAGAGCTGAGGAGAAATGCTTGAACCAAGAGCAGAGGCTGCA 22691
QY 350 GTGAACAGTGAACCCAGCTGATTTCCAGCTGGAAGACAGAGGAACTCTTTCCA 409
DB 22690 GTGAACCGAGATCATGCCACTGCACTCGAGCAACAGATGAGACTCATCTCAA 22631
QY 410 AAAAAAAAAAAAAAAAAAATGCAAGAAAGACATCATTAATCTTGAAGCTGGAGCATTAAT 469
DB 22630 AATTAATAAATAAATAAATAA---AAATTTAGCTGGAGATGAGACACA 22584
QY 470 TTTATGTGATGAATAATTCACATCTTTTGAAGAAATTAAGCATTTCTGATTAATGTAAT 529
DB 22583 TGTGTAGTCCCAATTCATAAAGGCTGAGGAGAGAAATCTTGAAGCAGAAATTTG 22524
QY 530 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 589
DB 22523 AGGCTGAGTGAAGTGAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 22464
QY 590 CTCTCAGTTGTCAGTCAAACTTTAATGATCTTTGGCCGGGTCGGTGGCTCAGCCTGT 649
DB 22463 CTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 22416
QY 650 AATCCAGCATTTTGGAGGCGGAGGCGGAGTGAATCAAGGTTAGAGATGAGACCAT 709
DB 22415 AATCCAGCATTTTGGAGGCGGAGGCGGAGTGAATCAAGGTTAGAGATGAGACCAT 22356
QY 710 CCTGGCTAACAGCGTGAACCTCGCTCTACTAATAAATAAATAAATTAAGCCGGTGGG 769
```


AC ADI12698;
 XX 22-APR-2004 (first entry)
 XX Human LPIN2 genomic DNA as given by ENSEMBL (Fig 3).
 DE de; type 2 diabetes; insulin resistance; human; lipin 2; LPIN2;
 XX linkage disequilibrium polymorphism; antidiabetic; gene therapy.
 XX Homo sapiens.
 OS MO2004001071-A2.
 XX 31-DEC-2003.
 XX 25-JUN-2003; 2003WO-GB002730.
 XX 25-JUN-2002; 2002GB-00014682.
 XX (OXAG-) OXAGEN LTD.
 XX Pullen J, Holdstock J;
 XX WPI; 2004-082513/08.
 XX Determining whether an individual is predisposed to type 2 diabetes
 PT and/or insulin resistance, useful for treating and/or preventing such
 PT disease, comprises typing the LPIN2 gene region or LPIN2 protein of the
 PT individual.
 XX Disclosure; Fig 3; 152pp; English.
 XX This invention relates to a novel method for determining whether an
 CC individual is predisposed to type 2 diabetes and/or insulin resistance.
 CC Specifically, it comprises typing the human lipin 2 (LPIN2) gene in order
 CC to detect at least one of the four recognised single nucleotide
 CC polymorphisms (SNPs) known to be associated with type 2 diabetes and
 CC insulin resistance. The present invention further describes detecting
 CC linkage disequilibrium polymorphisms that indicate a susceptibility or
 CC genetic predisposition to these conditions. The method comprises
 CC contacting a test agent with an LPIN2 mutated polymolecule or
 CC polypeptide and determining whether it is capable of binding and/or
 CC modulating the activity or expression of the molecule. Accordingly, these
 CC compositions exhibit antidiabetic activity and can be used to treat type
 CC 2 diabetes and/or insulin resistance using gene therapy. This
 CC polymolecule sequence is a human LPIN2 genomic DNA sequence of the
 CC invention.
 XX Sequence 97955 BP; 25989 A; 19721 C; 20487 G; 31058 T; 0 U; 700 Other;
 SQ
 Query Match 6.9%; Score 346.6; DB 12; Length 97955;
 Best Local Similarity 65.7%; Pred. No. 1.6e-47;
 Matches 610; Conservative 0; Mismatches 289; Indels 30; Gaps 6;

QY 329 GCGCAGAGGCTGAGGCTGACAGTGAACAGTATCACCCAGCTGATTCAGAGCTGAGAGA 388
 DB 15329 ATCCGGAGGTGAGGTGACATGAGCCGAGATCTGCGACCGACTCCAGCTTGAGTGA 15388
 QY 389 CAGAGGAGAGCCCTGTTCCAAAAAAGAAAAAATATGCAAG-----AAAGA 441
 DB 15389 GGGAGTGAACCTCTGCTTCAAAAAAAGAAAAAATATGCAAGCTTCCAAAAATGA 15448
 QY 442 CATCATAACTTGACCTGGGACATTAATTTATGATGAAATTCACAACTTTTAAAGAA 501
 DB 15449 AATGTACATATTATGATGATGTTGATTTGCGACCTCTATGACAGAAAGTTTACATTAC 15508
 QY 502 GAAATTAGATTTTCGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 561
 DB 15509 AATATATGCAAACTTGCGCCATTGGCAAGCCCTTTGTAAGAACAAATGATGCTTATTT 15568
 QY 562 AATATTTCTGAGAACTAGCTTCTCACTCTCTCAGTGTGAGTCAAAAACCTTAATGATC- 620
 DB 15569 GAGATGTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 15628
 QY 621 -----TTTGGCCGGGTGCGGTGCTCAGCCTGTAAATCCAGACTTTGGAGCCGAG 675
 DB 15629 AAGATATTGCTGAGGCTGCTCAGCCTGTAAATCCAGACTTTGGAGCCGAG 15688
 QY 676 CCGGTGATCACAAGTTTGAAGATGAGACCACTCTGCTTAACAGCTGAAACCTCTCTC 735
 DB 15689 TGGGTGATCAGAGCTGAGAGTGAACCATCTGCTTAACAGCTGAAACCTCTCTC 15748
 QY 736 TCTACTAAAAATACAAAAAATAGCGGGGTGGGTGCGGAGCCCTGTAGTCCAGCTGC 795
 DB 15749 TCTACTAAAAATACAAAAAATAGCGGGGTGGGTGCGGAGCCCTGTAGTCCAGCTGC 15808
 QY 796 TCAGAGGCTGAGGAGAGAAATGATGTAACCCGAGAGCCGAGCTTGCAGTGAAGCCGA 855
 DB 15809 TCGGAGGCTGAGGAGAGAAATGATGTAACCCGAGAGCCGAGCTTGCAGTGAAGCCGA 15868
 QY 856 GATTGGCCCACTGCACTCTCAGCTGGGCGACAGTGCAGACTCTGTCTTAAAAAAGAAAA 915
 DB 15869 GATTGGCCCACTGCACTCTCAGCTGGGCGACAGTGCAGACTCTGTCTTAAAAAAGAAAA 15928
 QY 916 AAAAAAAGTGAATGCTTGAAGCCA 944
 DB 15929 AAAAAAAGTGAATGCTTGAAGCCA 15957

RESULT 9
 ABQ75562
 ID ABQ75562 standard; DNA; 188888 BP.
 AC ABQ75562;
 XX
 DT 11-NOV-2002 (first entry)
 XX
 DE Human related CYP 27C1 clone RP11-30F3 SEQ ID NO:21.
 XX
 XX Cloning; characterisation; human; cytochrome P450; CYP 27C1; cytostatic;
 XX therapeutic; antidiabetic; antihypertensive; tuberculostatic; osteoporotic;
 XX dermatologic; antileptemic; gene therapy; vaccine; Vitamin D; diabetes;
 XX vitamin D metabolite deficiency; hyperparathyroidism; hypoparathyroidism;
 XX medullary carcinoma; psoriasis; sarcoidosis; tuberculosis; osteomalacia;
 XX chronic renal disease; vitamin D dependent rickets; anticonvulsant;
 XX fibrogenesis imperfecta ossium; osteitis fibrosa cystica; osteoporosis;
 XX osteopenia; osteosclerosis; renal osteodystrophy; rickets; steatorrhoea;
 XX glucocorticoid antagonism; idiopathic hypercalcaemia; tropical sprue;
 XX malabsorption syndrome; cholesterol steroid; lipid metabolic disorder;
 XX gene; de.
 XX
 OS Homo sapiens.
 XX
 XX MO200264765-A2.
 XX
 XX 22-AUG-2002.

Sequence 59588 BP: 15931 A; 13628 C; 13537 G; 16492 T; 0 U; 0 Other;

Query Match 6.9%; Score 345.2; DB 10; Length 59588;
 Best Local Similarity 68.9%; Pred. No. 2.6e-47;
 Matches 554; Conservative 0; Mismatches 203; Indels 47; Gaps 4;

136 CCAGGCTAGTGGCCGCGACACCCCACTGACGATATTTGGAGGCTGAGTGAAGAAAT 195
 46072 CCAGGCTATGAGCTCTCCTCCTTAATCCCAAGTTGGAGGTCAGAGTGAAGAAAT 46013
 196 GCTTGAAGCCAGAGTGAAGCAAGCCTAGGCAATAGTGAAGCCTGTGTCTA-TAA 254
 46012 GCTTGAAGCCAGAGTGAAGCAAGCCTAGGCAATAGTGAAGCCTGTGTCTA-TAA 45953
 255 AAAATTAATTAGCTGTGTGTTCTTGGCAGAGCCTGAG--CTAGCTACTGGAAGCTGAG 312
 45952 AAAATTAATTAGCTGTGTGTTCTTGGCAGAGCCTGAG--CTAGCTACTGGAAGCTGAG 45893
 313 GTGGAGAGATCAG-TGAGCCGAGAGGCTGAGGCTGCACTGAACAGTATCAACCCAGCTG 371
 45892 GTGGAGAGATCAGTGAAGCCAGAGGCTGAGGCTGCACTGAACAGTATCAACCCAGCTG 45833
 372 GATTCAGGCTGGAAGACAGAGGAGACCCCTGTTTCCAAAAAATAAAAAAATAAAAT 431
 45832 TACTTCAGGCTGGAAGACAGAGGAGACCCCTGTTTCCAAAAAATAAAAAAATAAAAT 45773
 432 CAAGAAAAGACATCATTAACCTGAGGACATTAATTATGTGATGAATTCACAAAT 491
 45772 AGACAAGATGTGTGCTGCTGCAATTCAGACATTTGAGAGGCCAGAGGAAGATC 45713
 492 CTTTGAAGAAATTAATGATTTCTGATTAATAATTAATTAATTAATTAATTAATTC 551
 45712 TCTTAGGCAAGAGTTCAGACAGCCCTAGGCAACATAGCAAGACCCATTTCTTAA 45653
 552 AAATGAATTAATTAATTTCTGAGAACTAGCTTCTCACTCTCAAGTTTCAATCAAACT 611
 45652 AAATGAATTAATTAATTTCTGAGAACTAGCTTCTCACTCTCAAGTTTCAATCAAACT 45628
 612 TTAATGTCTTTGGCCGAGGCTGAGGCTCAAGCCTGTAATCCAGCACTTTGGAGGCC 671
 45627 TTAATGTCTTTGGCCGAGGCTGAGGCTCAAGCCTGTAATCCAGCACTTTGGAGGCC 45576
 672 GAGGCGGTGATTCACAAAGTGAAGATTCAGACCATCTCTGCTTAACCGGTGAACCT 731
 45575 GAGGCGGTGATTCACAAAGTGAAGATTCAGACCATCTCTGCTTAACCGGTGAACCT 45516
 732 GCTCTCTACTAAATAATTAAGCCGGGTGCGGTGCGAGAGCCTGTAGTCCAG 791
 45515 TGTCTCTACTAAATAATTAAGCCGGGTGCGGTGCGAGAGCCTGTAGTCCAG 45456
 792 CTGCTCAGAGGCTGAGGAGAGAAATGCTGAACCCGGAAGGCGGAGCTTGCAGTGA 851
 45455 CTACTCAGAGGCTGAGGAGAGAAATGCTGAACCCGGAAGGCGGAGCTTGCAGTGA 45396
 852 CCGAATTTGCGGCACTGCACTCCAGCTGAGGAGCAAGTCCGAGACTTGTCTCAAAAA 911
 45395 CCGAATTTGCGGCACTGCACTCCAGCTGAGGAGCAAGTCCGAGACTTGTCTCAAAAA 45336
 912 AAAAAAAAAAAGTTGAATGCTC 935
 45335 TAAAAAATTAATAATTAAGCC 45312

RESULT 11
 AD112699
 ID AD112699 standard; DNA; 95103 BP.
 AC AD112699;
 XX 22-APR-2004 (first entry)
 XX Human LPIN2 genomic DNA as given by Genome Browser (Fig 4).
 XX

de; type 2 diabetes; insulin resistance; human; lipin 2; LPIN2;
 linkage disequilibrium polymorphism; antidiabetic; gene therapy.

OS Homo sapiens.
 PN MO2004001071-A2.
 XX 31-DEC-2003.
 XX 25-JUN-2003; 2003WO-GB002730.
 XX 25-JUN-2002; 2002GB-00014682.
 XX (OXAG-) OXAGEN LTD.
 PA Pullen J, Holdstock J;
 DR WPI; 2004-082513/08.
 XX Determining whether an individual is predisposed to type 2 diabetes
 PT and/or insulin resistance, useful for treating and/or preventing such
 PT disease, comprises typing the LPIN2 gene region or LPIN2 protein of the
 PT individual.
 XX disclosure; Fig 4; 152bp; English.

This invention relates to a novel method for determining whether an
 CC individual is predisposed to type 2 diabetes and/or insulin resistance.
 CC Specifically, it comprises typing the human lipin 2 (LPIN2) gene in order
 CC to detect at least one of the four recognized single nucleotide
 CC polymorphisms (SNPs) known to be associated with type 2 diabetes and
 CC insulin resistance. The present invention further describes detecting
 CC linkage disequilibrium polymorphisms that indicate a susceptibility or
 CC genetic predisposition to these conditions. The method comprises
 CC contacting a test agent with an LPIN2 mutated polynucleotide or
 CC polypeptide and determining whether it is capable of binding and/or
 CC modulating the activity or expression of the molecule. Accordingly, these
 CC compositions exhibit antidiabetic activity and can be used to treat type
 CC 2 diabetes and/or insulin resistance using gene therapy. This
 CC polynucleotide sequence is a human LPIN2 genomic DNA sequence of the
 CC invention.

Sequence 95103 BP: 25304 A; 19358 C; 19864 G; 29657 T; 0 U; 800 Other;

Query Match 6.9%; Score 344; DB 12; Length 95103;
 Best Local Similarity 66.0%; Pred. No. 4.1e-47;
 Matches 602; Conservative 0; Mismatches 280; Indels 30; Gaps 6;

46 GAGGCAAGAGATCGCTGAGCTCAGAAATGGAACAACCTTACCTAATAGTGAAC 105
 38917 GAGGCAAGAGATCGCTGAGCTCAGAAATGGAACAACCTTACCTAATAGTGAAC 38976
 106 CTTCTCTGTCAATAATTAATAAGAAATTT-----TCCAGGATGTGGCGTGCACC 156
 38977 CTTCTCTGTCAATAATTAATAAGAAATTT-----TCCAGGATGTGGCGTGCACC 39036
 157 CCCAGTGCAGCTATTTGGAGGCTGAGTGAAGAAATGCTTGAAGCAGAGATTGAAG 216
 39037 TGTATCCAGTACTTTGGGACACCAAGGCGGCGGATCAGTGAAGTGAAGATTGAAG 39096
 217 ACAAGCTTGAAGCAATATGAGACCTGTGTCTTAATAAAT-----AATTAGCTGTT 271
 39097 ACAAGCTTGAAGCAATATGAGACCTGTGTCTTAATAAAT-----AATTAGCTGTT 39156
 272 GTCTTGACACAGGCTGTGA--GCTAGCTACTCGAAGACTGAGTGAAGTCACTGAGTCA 328
 39157 GTCTTGACACAGGCTGTGA--GCTAGCTACTCGAAGACTGAGTGAAGTCACTGAGTCA 39216
 329 GCCAGAGAGCTGAGCTGCAAGTGAACAGTATCAACCCAGCTGATTCAGCTGGAAGA 388
 39217 ATCCGGAAGTGAAGTGAAGTGAATGAGTGTGTCACCGCACTCAGGCTGGTGA 39276
 389 CAGAGGAGACCTGTCTTCCAAAAAATAAAAAAATAATGCAAG-----AAAAGA 441

Db 39277 GGAAGTGAAGTCTGTCTCAAAAAAAAAAAAAAAAAAATACAGCTCCAAAAATGA 39336
QY 442 CATCATTAACCTGACCTGGACATTACTTTATGTGATGAATTCACATCTTTTAGGA 501
Db 39337 AATGTACATATTTATGATGTGTTGATTCTGGACCTCTTATGAGAAATTACCATTA 39336
QY 502 GAAATAGCATTTCTGATTAATATGATTTAATTAATTAATTAATTAATTAATTAATTA 561
Db 39397 AATATAGCAAACTTGCTGCATTTGGCAAGGCTTTGTGAACCAAAATGATGTTATTT 39456
QY 562 AATATTTCTGAGAACTAGCTTCTCACTCTCTCAAGTTTCAGTCAAACTTTAATGTC- 620
Db 39457 GGAGATGTGGGATGTGGGCTGTCTGTATCTGGGATTTATATGATCTCAATGATAGATA 39516
QY 621 -----TTTGGCCGGGGGCGGTGGCTCAGCGCTGTATCCAGACATTTGGAGGCCGAG 675
Db 39517 AAGATATTTGGTCAGAGTGTGGCTCAGCGCTGTATCCAGACATTTGGAGGCCGAG 39576
QY 676 CCGGTGATTCACAAGGTTAGAGATCGAGACATCTCTGCTAACACAGTGAACCTGCTC 735
Db 39577 TGGGTGATTCATACGTCAGAGATCGAGACATCTCTGCTAACACAGTGAACCTGCTC 39636
QY 736 TCTACTAAAAATTCAAAAAATTTGCGGGTGGCTGCCAGACCTCTGTAGTCCCACTGCTG 795
Db 39637 TCTACTAAAAATTCAAAAAATTTGCGGGTGGCTGCCAGACCTCTGTAGTCCCACTGCTG 39696
QY 796 TCAGAGGCTGAGAGCGAGAGATGTGTGAACCGGAGGCGGAGCTTCAGTGAAGCCGA 855
Db 39697 TCGGAGGCTGAGAGCGAGAGATGTGTGAACCGGAGGCGGAGCTTCAGTGAAGCCGA 39756
QY 856 GATTGCGCCACTGCACTCCAGCTGGGCGACAGTGGGAGACTGTCTCAAAAAAAAAA 915
Db 39757 GATGGCGCCACTGCACTCCAGCTGGGCGACAGTGGGAGACTGTCTCTCAAAAAA 39816
QY 916 AAAAAAAAAATT 927
Db 39817 AAAAAAAAAGAT 39828

RESULT 12
AAK81194/c
ID AAK81194 standard; DNA; 9979 BP.
XX
XX AAK81194;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36006.
XX
XX Human immune/haematopoietic; immune/haematopoietic antigen; cancer;
XX
XX cytosolic; gene therapy; vaccine; metastasis; da.
XX
XX Homo sapiens.
XX
XX W0200157182-A2.
XX
XX
XX 09-AUG-2001.
XX
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX

PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 08-SEP-2000; 2000US-0231968P.
PR 12-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233397P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 29-SEP-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 02-OCT-2000; 2000US-0237041P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR

PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-024617P.
 PR 01-NOV-2000; 2000US-024617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX MPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 XX useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 36006; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the

CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54952 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 XX Sequence 9979 BP; 2496 A; 2022 C; 2532 G; 2929 T; 0 U; 0 Other;
 Query Match 6.9%; Score 343; DB 4; Length 9979;
 Best Local Similarity 67.2%; Pred. No. 5.5e-47;
 Matches 627; Conservative 0; Mismatches 280; Indels 26; Gaps 9;
 QY 13 CTGTGGCTCAGCGGTGTATTCACACATTTGGAGC-----AGAGGATCGCTGAG 66
 DB 9080 CAGTGGCTCAGACTGTATTCACACATTTGGAGGCGCAAGTGGTGTATCTATGAG 9021
 QY 67 CTCAGGAATTTGAGCAACGCTTACGTAATATGTAACCTGTCTGTACAAATATATA 126
 DB 9020 GTTCAGAGATTAGACCAAGCTTGGCCCAAGGGGTAAACCCATCTCTAATAAATAGA 8961
 QY 127 AAGAAATTTCCAGGATGGTGGCGTGACCCCAAGTGCAGCTAATTTGGAGGAGTGGT 186
 DB 8960 AGAATTAGCCCAAGCATGTGCTCACACCTTAATCCAGACATTTGGAGGCGGAGGC 8901
 QY 187 AGAGGAATCTTGAAGCCAGAGTTGAAGACAAAGCTTAGGCAATAGTGAAGCCCTGT 246
 DB 8900 AGGTG--AATCAGAGGTGAGAGATCAAGACATCTGTGTTAACAAGTGAACCCCGT 8843
 QY 247 GTCTATN-----AAAAATATTAGCTGTGTCTTGGACACAGGCTGAG--CTAGCT 297
 DB 8842 TCTACTAAAAATACAAAAAAATTTAGCCGGGATGTGACAGCCCTGTATCTCCAGCT 8783
 QY 298 ACTCGGAAGACTGAGTGGAGAGATCAC--TGAGCCAGAGGCTGAGGCTGAGTGAACA 356
 DB 8782 ACTTAGAGGCTTAGGCGAGAGATCACTTAACCTGAGAGGCAAGAGCTTGCAGGAGGC 8723
 QY 357 GTGATCACCCAGCTGAATTCAGCTGGAAGACAGAGGAGACCTGTGTTCCAAAAA 416
 DB 8722 GAGATCAGCCCACTACATCCAGCTGGGCGACAGAGGAGACTGTGACCAAAAAA 8663
 QY 417 AAAAAAAAAAAAAAAGCAATCATTAACCTGACCTGG---ACATACTTT 472
 DB 8662 AAAAAACAGAAAGCTGTATTGTAATTTGAACCAAGTTCACATCATATATATATCA 8603
 QY 473 ATGTATGAATTTCAATCTTTTAGAAGAAATTAGCTTCTGTAATATATATATATA 532
 DB 8602 AAACGTTTACTTCACTATGATCCAAAGTCAATCTTCAATTATATAGAGCCACACA 8543
 QY 533 ATTATATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 592
 DB 8542 AAGAAACAAAGGAGATCATTTGAGCAATGTA--TGCAAGACATTCACATTTGAT 8484
 QY 593 TGAATTGATGACAAATCTTAATGATCTTTGGCCGGGTGCGGTGCTGACCCCTGTAT 652
 DB 8483 TCAATTGCTCTTAAGAAATTTCTCTTCCTGGTGGGCGGTGCTGACCCCTGTAT 8424
 QY 653 CCCAGCATTTGGAGGCGGAGGCGGAGTGAATCAACAAGTTAGAGATGAGACCATCT 712
 DB 8423 CCCAGCATTTGGAGGCGGAGGCGGAGTGAATCAACAAGTTAGAGATGAGACCATCT 8364
 QY 713 GGCTAACCGGTGAACCTGCTCTCTACTAATAAATAAATAAATAAATAAATAAATAA 772
 DB 8363 GGCTAATACGGTGAAACATCTCTAATAAATAAATAAATAAATAAATAAATAAATAA 8305
 QY 773 CAGAGCCCTGTATGCTCCAGCTGCTTACAGAGGCTGAGGAGAGATGCTGTGA--CCG 830
 DB 8304 TGGGGCTGTATGCTCCAGCTACTCGGAGGCTGAGGAGAGATGCTGTGA--CCG 8245
 QY 831 GGAGGCGAGCTTGCAGAGGAGGAGGAGTGGCCCACTGCACTCCAGCTGGGCGACAGT 890
 DB 8244 GGAGGTGAGCTTGCAGAGGAGGAGTGGCCCACTGCACTCCAGCTGGGCGACAG 8185

Db 24669 CTACCTGGAGGCTGAGGCAAGAAATCCCTTGAACCCAGAGGCTGACAGTGG 24728
Qy 355 CAGTATCAACCCAGCTGATTCAGCCTGAAAGACAGAGGAGACCCCTTTTCAAAAA 414
Db 24729 CCGAATCAACACATGACCTCCAGCTGGGTGACAGACGAGATTCCTCTCAAAAAA 24788
Qy 415 AAAAAAAAAAATGCAAGAAAGACATCAATACTTGACCTGGACATACTTTAT 474
Db 24789 AAAAAAAAAA-----ATTAGCTAGGGGTGCTGACGATACCTGTAA 24830
Qy 475 GATGATGAATTCACATCTTTAGAGAAATAGACATTTGATTAATGATTTAT 534
Db 24831 CCCAGTACTTGGAGCTGAGGACATGAGACTA-----CTTGAACCCAGAGAG 24881
Qy 535 TATATTATTAATTAATCAATGGAATTAATTAATTTGAGAACTAGCTTCTCCTCTC 594
Db 24882 GAGGTTGAAGTGAATGATTAATGCGTCATTGCGTTCCGGCTTGGGGACAGATGAGACT 24941
Qy 595 AGTTGCTACATAAACTTTAATGCTTTTGGCCGGGTGCGGTGCTCAGCCTGTATCC 654
Db 24942 CTGCTCAAAAAAATTAATAAATAAAGCCGGCGCAGTGGCTCAGCCTGTATCC 25001
Qy 655 GAGCACTTTGGAGGCGGAGGCGGCTGATCAAGGTTAGGAGATGAGACCATCTCG 714
Db 25002 CAGCACTTTGGAGGCGGAGGCGGAGATCAAGAGTCAAGAGATGAGACCATCTCG 25061
Qy 715 CTAAACGCTGAACCTGCTCTCTAATAAATAATACAAAAATTAAGCGGGTGGCGCA 774
Db 25062 CTAAACGCTGAACCTGCTCTCTAATAAATAATACAAAAATTAAGCGGGTGGCGCA 25121
Qy 775 GAGCCTGTATGTCCTCAGCTGCTCAGAGGCTGAGGAGAGAAATGTTGAAACCGGAG 834
Db 25122 GCGGCTGTATGTCCTCAGCTGCTCAGAGGCTGAGGAGAGAAATGCGTGAACCGGAG 25181
Qy 835 GGGAGCTTGCAGTGAAGCGGAGTGGCGCAGCTGCACTCCAGCTGGGAGACGTCCGAG 894
Db 25182 GCGGAGCTTGCAGTGAAGCGGAGTGGCGCAGCTGCACTCCAGCTGGGAGACGTCCGAG 25241
Qy 895 ACTGTGTCAAAAAAAAAAAAAAAAAA 924
Db 25242 TCCGTCTCAAAAAAAAAAAAAAAAAA 25271

RESULT 14
AB14581/C
ID AB14581 standard; DNA; 8288 BP.
XX
AC AB14581;
XX
DT 23-JAN-2002 (first entry)
XX
XX
DE Human nervous system related polynucleotide SEQ ID NO 6912.
XX
XX Human; nootropic; neuroprotective; cytosolic; dermatological; virucide;
KM immunosuppressive; antihistaminic; anti-HIV; antibacterial; vulnery;
KM antiparkinsonian; antidiabetic; antianemic; antitachycardic; cancer;
KM antirheumatic; hepatocytic; cerebroprotective; antiinflammatory;
KM antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KM antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KM neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186530P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-APR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-020515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225470P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232063P.
PR 14-SEP-2000; 2000US-0232064P.
PR 14-SEP-2000; 2000US-0232065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.

Db 4955 TGAACCGGAGGCGGAGCTTGCACTGACCGGAGATTGCCCACTGCTCCGCTGGG 4896
Qy 883 CGACAGTGGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAGTTGAAT 931
Db 4895 CCACAGAGGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAGAAAT 4847
RESULT 15
AAK85745/c
ID AAK85745 standard; DNA; 4515 BP.
XX AAK85745;
AC AAK85745;
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40557.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; da.
XX Homo sapiens.
XX WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-018650P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214686P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-021680P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226379P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228927P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235634P.
PR 27-SEP-2000; 2000US-0235635P.
PR 29-SEP-2000; 2000US-0235935P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237032P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.

PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251899P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI, 2001-483426/52.
 DR
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 40557; 3071bp + Sequence Listing; English.

CC AAK5951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent
 CC diagnosis and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 XX
 SQ Sequence 4515 BP; 1075 A; 1192 C; 1099 G; 1145 T; 0 U; 4 Other;

Query Match 6.8%; Score 338.6; DB 4; Length 4515;
 Best Local Similarity 67.1%; Pred. No. 2.8e-46;
 Matches 634; Conservative 0; Mismatches 239; Indels 72; Gaps 8;

QY 15 GTGGCTACCGCGTATTCCTCCAGACCTTTGGG-----AGCAGAGAGATCGCTTGAGC 67
 DB 941 GTGGCTATCCCTGTATCCAGACCTTTGGGAGGCCAAGCGGGAGATCACTGAGC 882
 QY 68 TCAGGAATTGAGACAGACCTACAGTAACTGTCTGTGTAACAATAATAA 127
 DB 881 CCAGAAATTCAAGACAGCTGGGCAACATAGTAGACCTGTCTCCAAAAA 822
 QY 128 AGAATTTTCA-----GGCATGTGGCGTGCACCCCAAGTCCAGCTATT 173
 DB 821 AAAAAAACCAGAAAGAGTGGGTGGCTCACCCCTGTATCCAGACCTTT 762
 QY 174 GGGAGGCTGAGGTAGAGAGATGTTGAAGCCAGAGATTAAGACAAGCTTAGCAAT 233
 DB 761 GGGAGGCGGAGGTGGTGA--TCACGAGGTCAAGAGATTGAGCAATCAGGTTACAG 704
 QY 234 AGTGAGACCTGTGTCT-----ATAAAAAATAATTAGCTGTGTGTTGGCACA 282

DB 703 GGTGAACCCCATCTCTACTAATAAATACAAAAAATATAGCCAGGCGTGTGTGG 644
 QY 283 GGCCTGCAG--CTAGCTTCTCGAAAGTGAAGTGGAGATGAC--TAGCCCAAGAGC 339
 DB 643 CACCTGTAGTCCCACTACTTGGAGGCTGAGGCGAGAGAAATGGCGTGAACCCAGAGAGC 584
 QY 340 TAGAGCTGACAGAACAGTATACACCAAGCTGATTTCCAGCTGGAGACAGAGGAGAC 399
 DB 583 AAGGTTACAGAGACCGAGATAGCGCACTACCTCCAGCTGGGCAAGAGAGAC 524
 QY 400 CCTGTTCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 459
 DB 523 TCTGCTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 464
 QY 460 GGAATTAATTTATGTATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 519
 DB 463 GGGCGGTGGCTACCGCTGTATCCAGACACTTTGGAGGCTGA----- 420
 QY 520 AAAATGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 579
 DB 419 -----TGGGGTGAATCAAGGTGAGAGATGAGACATCTGGCTAACAGGT 369
 QY 580 GCTTCTCACTCTCTCAGTTGTCAGTCAAAACTTTAATGTCCTTTGGCCGGGTGCGTGGC 639
 DB 368 GAAACCCCGTCTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 319
 QY 640 TCACGCTGTATTCACAGACTTTGGAGGCGGAGGCGGTGATCAAGGTTAGAGA 699
 DB 318 TCACGCTGTATTCACAGACTTTGGAGGCGGAGGCGGTGATCAAGGTTAGAGA 259
 QY 700 TGAAGACATCTGCTAACAAGTGAACCTGCTCTACTAATAAATAAATAAATAAATAA 759
 DB 258 TGAAGACATCTGCTAACAAGTGAACCTGCTCTACTAATAAATAAATAAATAAATAA 199
 QY 760 CCGGTCGCTGTCAGACGCTGTATGTCACAGCTGCTCAAGAGGCTGAGGCGAGAGATG 819
 DB 198 CCGGTCGCTGTCAGAGGCGCTGTATGTCACAGCTGCTCAAGAGGCTGAGGCGAGAGATG 139
 QY 820 GTGTGAACCCGGAGGCGAGGCTTGCATGAGCCGAGATTTGGCCACTGCCAGCCT 879
 DB 138 GGTGAACCCGGAGGCGAGGCTTGCATGAGCCGAGATTTGGCCACTGCCAGCCT 79
 QY 880 GGGCGCAGGTGAGACCTGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 924
 DB 78 GGGCGCAGGTGAGACCTGTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 34

Search completed: December 26, 2004, 18:48:42
 Job time : 1456.16 secs

Db 20240 ATGTGAAGCCCTCAGCCATCTGAAGGACAGTGTACAGCAATGATCAAAAAAGAAAA 20299
QY 301 CCACAGGCCCTTCCCTTCCCATCTTGAATGAAGAGCTCTTCAATTTCCATAGTAGT 360
Db 20300 CCACAGGCCCTTCCCTTCCCATCTTGAATGAAGAGCTCTTCAATTTCCATAGTAGT 20359
QY 361 AAATTTTCTAGATACAGCTTTGAGAGCTCAAAAGTACTGAAAAGAAAGCTCCCATTAAG 420
Db 20360 AAATTTTCTAGATACAGCTTTGAGAGCTCAAAAGTACTGAAAAGAAAGCTCCCATTAAG 20419
QY 421 GAAATTTATCTTAAGATACGTAAATGATACATAATTTTGTACATTTGGAAATATATAGT 480
Db 20420 GAAATTTATCTTAAGATACGTAAATGATACATAATTTTGTACATTTGGAAATATATAGT 20479
QY 481 TGTAGCCTGGCGCGGTGCTCACGCTGTAAATCCAGCCCTTTGGAGGCCAGATGGG 540
Db 20480 TGTAGCCTGGCGCGGTGCTCACGCTGTAAATCCAGCCCTTTGGAGGCCAGATGGG 20539
QY 541 CAGATCTATGAGGTCAAGAGTGTGAGACCAAGCTAGCCAACTATGTGAAAACCCCGTCTTA 600
Db 20540 CAGATCTATGAGGTCAAGAGTGTGAGACCAAGCTAGCCAACTATGTGAAAACCCCGTCTTA 20599
QY 601 CTAAAGATCAAAAAAATTTAGCCAGGTGTGTGGCGCACACTGTAAACCCAGCTGTCCA 660
Db 20600 CTAAAGATCAAAAAAATTTAGCCAGGTGTGTGGCGCACACTGTAAACCCAGCTGTCCA 20659
QY 661 GAGAGTGAAGCAGAGAAATTTGTAACCCAGAGGCCAGAGGTGCAGCCAGCAAGATCA 720
Db 20660 GAGAGTGAAGCAGAGAAATTTGTAACCCAGAGGCCAGAGGTGCAGCCAGCAAGATCA 20719
QY 721 CACCAATGCACTGTACCTGAGATGACAGGCCAAAGATCCCACTCAAAAAAATTTAAAAA 780
Db 20720 CACCAATGCACTGTACCTGAGATGACAGGCCAAAGATCCCACTCAAAAAAATTTAAAAA 20779
QY 781 AAGAAATATGTAAAGTTGTCTATAACAATAATGAAGCAGTGAAGCAAAAGTCTAAAG 840
Db 20780 AAGAAATATGTAAAGTTGTCTATAACAATAATGAAGCAGTGAAGCAAAAGTCTAAAG 20839
QY 841 CCTATGACCATGTGTAATAGAAATCTGTGGGAAACATATATAGGAACTTAACCCAGT 900
Db 20840 CCTATGACCATGTGTAATAGAAATCTGTGGGAAACATATATAGGAACTTAACCCAGT 20899
QY 901 CCTGGAAGTAAGTTTGGAAAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTTGAAAA 960
Db 20900 CCTGGAAGTAAGTTTGGAAAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTTGAAAA 20959
QY 961 AAAAAATTAATAATCAGTTTACGTGTGTGGAGAAATGGGATAGGGACTTAATAGAGAAAT 1020
Db 20960 AAAAAATTAATAATCAGTTTACGTGTGTGGAGAAATGGGATAGGGACTTAATAGAGAAAT 21019
QY 1021 CAAATAGGAATGTTTCATGTGATGTTTAAAGACCCTGTGTAAAGGTGAAGCAATTAATTA 1080
Db 21020 CAAATAGGAATGTTTCATGTGATGTTTAAAGACCCTGTGTAAAGGTGAAGCAATTAATTA 21079
QY 1081 TCTGGCAACATCCGGGGAACTTTTATTAAGGTAAATGCTTGGCAATTTAAATAGAGACA 1140
Db 21080 TCTGGCAACATCCGGGGAACTTTTATTAAGGTAAATGCTTGGCAATTTAAATAGAGACA 21139
QY 1141 GAGAAATGTAGACAGTTGATGAGTCAAGATTGAAATTTCTGCACAGACATGTGAAAAGAA 1200
Db 21140 GAGAAATGTAGACAGTTGATGAGTCAAGATTGAAATTTCTGCACAGACATGTGAAAAGAA 21199
QY 1201 AGACAGGTAGGCAAGAGAGTTTGAAGAGATTATCAAGACAGAAATTAATGTGCTGCCAGT 1260
Db 21200 AGACAGGTAGGCAAGAGAGTTTGAAGAGATTATCAAGACAGAAATTAATGTGCTGCCAGT 21259
QY 1261 GGCATCTAGCTGATCTATCTGAGGGAAGAAAGTAAGCAAGCTTGTGCTGATAGT 1320
Db 21260 GGCATCTAGCTGATCTATCTGAGGGAAGAAAGTAAGCAAGCTTGTGCTGATAGT 21319
QY 1321 TATGAAGAGAGTGAAGGCTTCAAGACCTACAGGTGTTGATTAATAGAAATGATTTG 1380

Db 21320 TATGAAGAGAGTGAAGGCTTCAAGACCTACAGGTGTTGATTAATAAGAAATGATTTG 21379
QY 1381 GAGAAAGATTAATCTGTGAGAGTGAATTTTCAAGCTTTGAGTGACTCTCAATATCCAGA 1440
Db 21380 GAGAAAGATTAATCTGTGAGAGTGAATTTTCAAGCTTTGAGTGACTCTCAATATCCAGA 21439
QY 1441 CACTGTGCTAAATGCTTCAAGAGACATGATCCCTGCCCTCAAGGGAGCTTACAGCCAAAAAC 1500
Db 21440 CACTGTGCTAAATGCTTCAAGAGACATGATCCCTGCCCTCAAGGGAGCTTACAGCCAAAAAC 21499
QY 1501 AAGAGATTAAGAAATACACCAATATCTATTATAGGACACTTGTGTAGAAATATCAAGAAAG 1560
Db 21500 AAGAGATTAAGAAATACACCAATATCTATTATAGGACACTTGTGTAGAAATATCAAGAAAG 21559
QY 1561 AAATACGATCTTAGTACTGTAGATGTGCAACGGCATCAAAAGATATCTTCTAGTTTCAAGAA 1620
Db 21560 AAATACGATCTTAGTACTGTAGATGTGCAACGGCATCAAAAGATATCTTCTAGTTTCAAGAA 21619
QY 1621 GTTTCAGATCCGCCCGGGCGGGTGGCTCACGCTGTAAATCCAGACATTTTGGAGGCCGA 1680
Db 21620 GTTTCAGATCCGCCCGGGCGGGTGGCTCACGCTGTAAATCCAGACATTTTGGAGGCCGA 21679
QY 1681 GGCGGGTGATCACAAGGTCAAGAGATCAAGACATCTGTGTAAACACGGTGAACCCCG 1740
Db 21680 GGCGGGTGATCACAAGGTCAAGAGATCAAGACATCTGTGTAAACACGGTGAACCCCG 21739
QY 1741 TCTCTACAAAAATATTAATAAATTTAGCCAGGGTGTGTGGCGGCTCTAGTCCAGCT 1800
Db 21740 TCTCTACAAAAATATTAATAAATTTAGCCAGGGTGTGTGGCGGCTCTAGTCCAGCT 21799
QY 1801 ACTCAGAGAGCTGAGGACAGAGAAATGCGGTGAACCCGGAGGTAGAGTTTGCCTGAGCCG 1860
Db 21800 ACTCAGAGAGCTGAGGACAGAGAAATGCGGTGAACCCGGAGGTAGAGTTTGCCTGAGCCG 21859
QY 1861 AGATGCGCGCACTGCGCTTCAGACTGTGGCGACAGAGTGAAGCTGCGTCTCAAAAAAATA 1920
Db 21860 AGATGCGCGCACTGCGCTTCAGACTGTGGCGACAGAGTGAAGCTGCGTCTCAAAAAAATA 21919
QY 1921 AAAAAAATAAATTTTCAATCTTAACACATGCAATTTCAACAGCTAGAAATAGAG 1980
Db 21920 AAAAAAATAAATTTTCAATCTTAACACATGCAATTTCAACAGCTAGAAATAGAG 21979
QY 1981 AGCATGTTACAGGAGAGAAAAATGTTTTCAGCAAAAGTACAGAGTAGGGAAATAGAGAT 2040
Db 21980 AGCATGTTACAGGAGAGAAAAATGTTTTCAGCAAAAGTACAGAGTAGGGAAATAGAGAT 22039
QY 2041 ATGTTCAAGAAAGAGACCCAGAGTCATGTTTGTAGGGTTTAAAGAAACACAGTGT 2100
Db 22040 ATGTTCAAGAAAGAGACCCAGAGTCATGTTTGTAGGGTTTAAAGAAACACAGTGT 22099
QY 2101 TTGCAATCTCCAGGTTCCATTAATGTCGTTAATGAATCAATAGTGTGTAAGCAACCTGCA 2160
Db 22100 TTGCAATCTCCAGGTTCCATTAATGTCGTTAATGAATCAATAGTGTGTAAGCAACCTGCA 22159
QY 2161 TTTTAAAAAATGAATAATATGATGAGAGAAATGAATAATTTAGCATGATTAACATT 2220
Db 22160 TTTTAAAAAATGAATAATATGATGAGAGAAATGAATAATTTAGCATGATTAACATT 22219
QY 2221 TTGAAAGAGCAAGTATATTTTCTGCAACTTTTGTCTCAATTTGTAATCTGATTAATTT 2280
Db 22220 TTGAAAGAGCAAGTATATTTTCTGCAACTTTTGTCTCAATTTGTAATCTGATTAATTT 22279
QY 2281 TTATGTATGATGTGAATTCAGATACATATTAATTTCTTACTAGTACAGATCAAAA 2340
Db 22280 TTATGTATGATGTGAATTCAGATACATATTAATTTCTTACTAGTACAGATCAAAA 22339
QY 2341 AATCTTAAAGCAGCTGCTGCTTAATCTTATTTTTCAGAGAGAAATCCAGATC 2400
Db 22340 AATCTTAAAGCAGCTGCTGCTTAATCTTATTTTTCAGAGAGAAATCCAGATC 22399
QY 2401 TGAAGAGCAAAACATTTTGTCTGAGGTTATAGAACCAAGTTATGCCATTGCTTAAAGTGA 2460
Db 22400 TGAAGAGCAAAACATTTTGTCTGAGGTTATAGAACCAAGTTATGCCATTGCTTAAAGTGA 22459

QY 2461 TTCTTAGTAAATTTCTTCCCACTAGTGCATCTGCACTTCTAGTCTGTGGCCTGA 2520
DB 22460 TTCTTAGTAAATTTCTTCCCACTAGTGCATCTGCACTTCTAGTCTGTGGCCTGA 22519
QY 2521 AATPACGAATATATATAGTGAACAGATTCACAAAGCTGGGGAAATATATGGGTAGTG 2580
DB 22520 AATPACGAATATATATAGTGAACAGATTCACAAAGCTGGGGAAATATATGGGTAGTG 22579
QY 2581 GCTGAGAGCCTCATTTTCTAAGAAATGTGACCTTAGCAGGGATGTGGCTCACACCT 2640
DB 22580 GCTGAGAGCCTCATTTTCTAAGAAATGTGACCTTAGCAGGGATGTGGCTCACACCT 22639
QY 2641 ATATATCCAGCACTTTGGAGGCCAAGTCAAGAAATCGCTTGAACCCAAAGATTCAAG 2700
DB 22640 ATATATCCAGCACTTTGGAGGCCAAGTCAAGAAATCGCTTGAACCCAAAGATTCAAG 22699
QY 2701 CTAGCATGGGCAACATAGCAAGACCTCATCTCTACAAAAAATTTAAATTAATCAGCTAGCA 2760
DB 22700 CTAGCATGGGCAACATAGCAAGACCTCATCTCTACAAAAAATTTAAATTAATCAGCTAGCA 22759
QY 2761 TGTGGCATACGCTGTAGTCCCACTACCTGGAAAGTAGTGGTGAATCGCTTGACA 2820
DB 22760 TGTGGCATACGCTGTAGTCCCACTACCTGGAAAGTAGTGGTGAATCGCTTGACA 22819
QY 2821 CAGGAGTTGAGGCTTAAGGTAGGCAATGATCACACAATGCACTCCAGCTTGAGTACAG 2880
DB 22820 CAGGAGTTGAGGCTTAAGGTAGGCAATGATCACACAATGCACTCCAGCTTGAGTACAG 22879
QY 2881 AGGAAGACCCCTGCTCCATAAAAAAGAAATGTGATTTTATTCCTTAGACAGTACAGT 2940
DB 22880 AGGAAGACCCCTGCTCCATAAAAAAGAAATGTGATTTTATTCCTTAGACAGTACAGT 22939
QY 2941 CATTAAGTCAATTAAGTTGAGTTGAGAGAAATATATGATCAGAGAAATTTATATCACT 3000
DB 22940 CATTAAGTCAATTAAGTTGAGTTGAGAGAAATATATGATCAGAGAAATTTATATCACT 22999
QY 3001 GTGGTCTGATGATATATGAAGAAATTAAGAGATGAGTCAAGGATTCACATTAAAGTG 3060
DB 23000 GTGGTCTGATGATATATGAAGAAATTAAGAGATGAGTCAAGGATTCACATTAAAGTG 23059
QY 3061 TTGTGTTGTTGTTTGAAGCAGAGTCTCTTTTGTACCCAGGCTAGTGAAGTGAGTG 3120
DB 23060 TTGTGTTGTTGTTTGAAGCAGAGTCTCTTTTGTACCCAGGCTAGTGAAGTGAGTG 23119
QY 3121 CAGTCAATGCTCACCGCAGCTCAAACTCCAGGCTCAAAATTTCTTCCAGCTCGGCT 3180
DB 23120 CAGTCAATGCTCACCGCAGCTCAAACTCCAGGCTCAAAATTTCTTCCAGCTCGGCT 23179
QY 3181 CCCAAAGTCTGGAATTAACAGTGTGAGGCCAAAGGTTTATGATGTGTCTGGCCTAGT 3240
DB 23180 CCCAAAGTCTGGAATTAACAGTGTGAGGCCAAAGGTTTATGATGTGTCTGGCCTAGT 23239
QY 3241 GCTCTCAAAATTCATGAGCAGACAAAGTGAACCGGGAACCTGACTCAACAAAGCTGGGT 3300
DB 23240 GCTCTCAAAATTCATGAGCAGACAAAGTGAACCGGGAACCTGACTCAACAAAGCTGGGT 23299
QY 3301 TAAGCTGAGCCTCTGATTTCTAATCATGAGTCAAGCTGATGAGTGTGCTGTCAAG 3360
DB 23300 TAAGCTGAGCCTCTGATTTCTAATCATGAGTCAAGCTGATGAGTGTGCTGTCAAG 23359
QY 3361 GCCAAGCACTGAGCAGCAAGAGTCTAGTGAATTAATGATGATCAAGGTGATTTATGCT 3420
DB 23360 GCCAAGCACTGAGCAGCAAGAGTCTAGTGAATTAATGATGATCAAGGTGATTTATGCT 23419
QY 3421 AGTGAATATAGAAATGGAAGAAATGTAATAATCAGTAACAAAGAGAGTTCACCTTG 3480
DB 23420 AGTGAATATAGAAATGGAAGAAATGTAATAATCAGTAACAAAGAGAGTTCACCTTG 23479
QY 3481 GTAATGTGAGCATGAGAGGAAAGATGGGGCAAAATCTGGTTTGTGTTGACT 3540
DB 23480 GTAATGTGAGCATGAGAGGAAAGATGGGGCAAAATCTGGTTTGTGTTGACT 23539

QY 3541 GACGAGGAATTTAGTCTCTATTAACAGAAATPAGSAGAAAGTTGTTGGAGAAA 3600
DB 23540 GACGAGGAATTTAGTCTCTATTAACAGAAATPAGSAGAAAGTTGTTGGAGAAA 23599
QY 3601 GAGAGTCTGTTTCAGACGTGTGAGGTCCAGGTGAGACAGATCTCCAAAGGAAAT 3660
DB 23600 GAGAGTCTGTTTCAGACGTGTGAGGTCCAGGTGAGACAGATCTCCAAAGGAAAT 23659
QY 3661 GAGCAGTAGGCAACCTAAAGAAATCTGTGCTCAGAGGAGAGCTGTGAGCTGT 3720
DB 23660 GAGCAGTAGGCAACCTAAAGAAATCTGTGCTCAGAGGAGAGCTGTGAGCTGT 23719
QY 3721 AGATCTGAGGTCATCAGACATAGAGTTTAAAGCAAGSAGTGGCAACCAAGAGAGC 3780
DB 23720 AGATCTGAGGTCATCAGACATAGAGTTTAAAGCAAGSAGTGGCAACCAAGAGAGC 23779
QY 3781 AAATACACAAAGAGAGAGAGTATGATGAGACCTTTGCTTTTATAGATGAGAGAG 3840
DB 23780 AAATACACAAAGAGAGAGAGTATGATGAGACCTTTGCTTTTATAGATGAGAGAG 23839
QY 3841 ACAGGAATGAAGAAATGAAGGAGAGCAGCTGTGAGAAATGTAAGCATCTGAAAAA 3900
DB 23840 ACAGGAATGAAGAAATGAAGGAGAGCAGCTGTGAGAAATGTAAGCATCTGAAAAA 23899
QY 3901 ATACACATGATGAGAGTCAAGGGAAGAAATTTCAAGSAGAGGATATGGTGAACA 3960
DB 23900 ATACACATGATGAGAGTCAAGGGAAGAAATTTCAAGSAGAGGATATGGTGAACA 23959
QY 3961 GTATTAACAGCATACAGAAATACAGCTAAAGTCAATCTGCTGATGATGACCTTGTG 4020
DB 23960 GTATTAACAGCATACAGAAATACAGCTAAAGTCAATCTGCTGATGATGACCTTGTG 24019
QY 4021 ATTTGTAGGGAACAATAATTAAGAAATTTATGTTGGATATAGAGGCAAAAG 4080
DB 24020 ATTTGTAGGGAACAATAATTAAGAAATTTATGTTGGATATAGAGGCAAAAG 24079
QY 4081 GAAGAGTTATCCAAATCAACAGGTGGAGTGAAGGATGAGTCTCCAAAGGTGAGGCA 4140
DB 24080 GAAGAGTTATCCAAATCAACAGGTGGAGTGAAGGATGAGTCTCCAAAGGTGAGGCA 24139
QY 4141 TCAGTAATGTGGGAAGGGGCAAGAGCATTCATGCGCATCCAGGCAAGCCCTCCA 4200
DB 24140 TCAGTAATGTGGGAAGGGGCAAGAGCATTCATGCGCATCCAGGCAAGCCCTCCA 24199
QY 4201 GAAGCTCATGAGAGTTCAGCTATCCAGAGGTCTGTGACCTTAATCTTTCTGGGTT 4260
DB 24200 GAAGCTCATGAGAGTTCAGCTATCCAGAGGTCTGTGACCTTAATCTTTCTGGGTT 24259
QY 4261 TGCATAGGCTTCAATGTGTGAGGCAATGATTTAAATCTATGAGGCACTGGTATCAACT 4320
DB 24260 TGCATAGGCTTCAATGTGTGAGGCAATGATTTAAATCTATGAGGCACTGGTATCAACT 24319
QY 4321 AACCTTCAACCCCTCCCTCCCTCAATCATGCTGTGCTTTCCAGTACAGGCTCCTA 4380
DB 24320 AACCTTCAACCCCTCCCTCCCTCAATCATGCTGTGCTTTCCAGTACAGGCTCCTA 24379
QY 4381 TCCATAAGCTAACCAATGTGTGCGCAGTATCAGTCAACTCTACAAAAAGACATCACTTTG 4440
DB 24380 TCCATAAGCTAACCAATGTGTGCGCAGTATCAGTCAACTCTACAAAAAGACATCACTTTG 24439
QY 4441 GAGATTTCAAGATTTTGAAGTGGCTGTCAAGAAATTTAGTTGAAGATCAAAATATAT 4500
DB 24440 GAGATTTCAAGATTTTGAAGTGGCTGTCAAGAAATTTAGTTGAAGATCAAAATATAT 24499
QY 4501 TTCACATATCAAGTGTGCTATTTTATATCAGGCGCATTAATGTGTTTAAACAAAG 4560
DB 24500 TTCACATATCAAGTGTGCTATTTTATATCAGGCGCATTAATGTGTTTAAACAAAG 24559
QY 4561 AGGTGATTAATTCAGATTTTCTTTTAAAGCTTACATGATGACAGTGTGATGATG 4620
DB 24560 AGGTGATTAATTCAGATTTTCTTTTAAAGCTTACATGATGACAGTGTGATGATG 24619
QY 4621 ATTTGGATGAGGCAATCTTTTTTTTGAAGATGTTATTCCTCCGCACTTCTC 4680

[illegible]

Db	25700	CAADATATTGATTTTAGTAGAAGTGAAGCTATGAAAAAGAACAGCCAGAGGGGCAAG	257575
Qy	5761	AAAAATTAGAGATTTTGTGGAACCAAGAAGAGAGTCCTCAGGAAGAAGCATGTCT	5820
Db	25760	AAAAATTAGAGATTTGTGGAACCAAGAAGAGAGTCCTCAGGAAGAAGCATGTCT	25819
Qy	5821	ATGATGCCAAATCTCTCAAAAGATTAGAAATTAAGAGTATCCATTGGGTTTCATAGAAAA	5880
Db	25820	ATGATGCCAAATCTCTCAAAAGATTAGAAATTAAGAGTATCCATTGGGTTTCATAGAAAA	258797
Qy	5881	GTCATGGGAAACCATGTGTAAAAAACATTGTGATGATGACAAATCGTTGCAAAAGCATTTT	5940
Db	25880	GTCATGGGAAACCATGTGTAAAAAACATTGTGATGATGACAAATCGTTGCAAAAGCATTTT	259398
Qy	5941	TATAGGGGGATGGAATTTTGTATTTTCAGAGACAAACAGTTCATACATGCGAAAGATCTA	6000
Db	25940	TATAGGGGGATGGAATTTTGTATTTTCAGAGACAAACAGTTCATACATGCGAAAGATCTA	259989
Qy	6001	GTGGTGACCAACGGAGATTAGTGTCTGAAAGTGATTTGGAAGACAGATCATTTGAGCTGA	6060
Db	26000	GTGGTGACCAACGGAGATTAGTGTCTGAAAGTGATTTGGAAGACAGATCATTTGAGCTGA	26055
Qy	6061	GGTGGGTAGAGTGTCTTCATGGAACATTAATGTCATGGAATCAACAGCTGTGATCCAG	6120
Db	26060	GGTGGGTAGAGTGTCTTCATGGAACATTAATGTCATGGAATCAACAGCTGTGATCCAG	26119
Qy	6121	TGCCACATCTTCAGTGAATGACAGAGAGGATTGAGATTCAGTGAATGACCGCTAAAA	6180
Db	26120	TGCCACATCTTCAGTGAATGACAGAGAGGATTGAGATTCAGTGAATGACCGCTAAAA	261797
Qy	6181	GAAGATTAATGGAAGATGTGG	6201
Db	26180	GAAGATTAATGGAAGATGTGG	26200
RESULT 2			
AX571875			
LOCUS	AX571875	53332 bp	DNA
DEFINITION	Sequence 3 from Patent WO02061060.	linear	PAT 29-NOV-2002
ACCESSION	AX571875		
VERSION	AX571875.1	GI:26004000	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS	1		
TITLE	Yan, C., Ketchum, K., di Francesco, V., and Beasley, E. M.		
JOURNAL	Isolated human kinase proteins, nucleic acid molecules encoding		
	human kinases proteins, and uses thereof		
	Patent: WO 02061060-A 3 08-AUG-2002;		
	PE Corporation (NY) (US)		
FEATURES			
source	1..53332		
	/location/Qualifiers		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	100.0%; Score 6201; DB 6; Length 53332;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 6201; Conservative	0; Mismatches	0; Indels	0; Gaps
Qy	1	CACCATGGGCTCAGCGCAATTTATGAAGCCAAATATACGCCCTTAAAAATGAAATGTGAC	60
Db	20000	CACCATGGGCTCAGCGCAATTTATGAAGCCAAATATACGCCCTTAAAAATGAAATGTGAC	200595
Qy	61	CTAAATACCCAGAGAAGACTCCCTTTGTGAAGATTTGTAACTAAAAATTAATATGAGTAGAG	120
Db	20060	CTAAATACCCAGAGAAGACTCCCTTTGTGAAGATTTGTAACTAAAAATTAATATGAGTAGAG	20119
Qy	121	TTAATAGTTCTAATGGAATGTGAACCCAGAGCCATATCGCGCTAGCAAAATGGCAGAA	180

Db 20120 TTAATAGTCTAATGGAATGCTGAAACCCAGAGCATATCAGGCTAGGCAAAATGGCACA 20179
Qy 181 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGTCTAAAGAAA 240
Db 20180 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTCAGCGCTTAATGATGTCTAAAGAAA 20239
Qy 241 ATGTAAAGCCCTCGACCATCTGAAAGCAGAGTTTACAGCAATTTATCAAAAGAAAA 300
Db 20240 ATGTAAAGCCCTCGACCATCTGAAAGCAGAGTTTACAGCAATTTATCAAAAGAAAA 20299
Qy 301 CCACAGGCGCTTCCCTTCCCTCACTTGAATGTAAGAGCTTCTCAATTTCCATAGT 360
Db 20300 CCACAGGCGCTTCCCTTCCCTCACTTGAATGTAAGAGCTTCTCAATTTCCATAGT 20359
Qy 361 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGAGTCTGAAAGAAAGCTCCCATTCAG 420
Db 20360 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGAGTCTGAAAGAAAGCTCCCATTCAG 20419
Qy 421 GAAATTTATCTTAAGTACTGTAAATGATCTAATTTTGTACATTTGGAAATATATAGT 480
Db 20420 GAAATTTATCTTAAGTACTGTAAATGATCTAATTTTGTACATTTGGAAATATATAGT 20479
Qy 481 TGTAGCTGCGCGGCTGCTCAGCGCTGTAATCCAGCGCTTGGAGGCGCAGAGTGG 540
Db 20480 TGTAGCTGCGCGGCTGCTCAGCGCTGTAATCCAGCGCTTGGAGGCGCAGAGTGG 20539
Qy 541 CAGATCATAGAGTCAAGAGTCTTGAAGCAGAGCTAGCCACATGCTGAAAACCCCGTCT 600
Db 20540 CAGATCATAGAGTCAAGAGTCTTGAAGCAGAGCTAGCCACATGCTGAAAACCCCGTCT 20599
Qy 601 CTAAAGATCAAAAAATTTAGCCAGGTGTGTGGCGGCAACCTGTAAACCCCATGCTG 660
Db 20600 CTAAAGATCAAAAAATTTAGCCAGGTGTGTGGCGGCAACCTGTAAACCCCATGCTG 20659
Qy 661 GAGAGTAGAGCAGAGAAATTTGCTTGAACCCAGAGCAGAGGTGCGAGCAGCAAGATCA 720
Db 20660 GAGAGTAGAGCAGAGAAATTTGCTTGAACCCAGAGCAGAGGTGCGAGCAGCAAGATCA 20719
Qy 721 CACCAATGCACTGTAGCTGTAGTACAGAGGCAAGACTCCAACTCAAAAAA 780
Db 20720 CACCAATGCACTGTAGCTGTAGTACAGAGGCAAGACTCCAACTCAAAAAA 20779
Qy 781 AAGAAATATGTAAGTGTGTCTAATCAAAATTAATAGCAGTGAAGCAAGTGTCTAAG 840
Db 20780 AAGAAATATGTAAGTGTGTCTAATCAAAATTAATAGCAGTGAAGCAAGTGTCTAAG 20839
Qy 841 CCTATGACCATGTAATCTAGAAATCTGTGGGAACATATTAAGGGAACCTTAACCAT 900
Db 20840 CCTATGACCATGTAATCTAGAAATCTGTGGGAACATATTAAGGGAACCTTAACCAT 20899
Qy 901 CCTGGAAGTAAAGTGTGGAAGAAATGTTTGAAGCAAAAGGTTAAAGAGTGA 960
Db 20900 CCTGGAAGTAAAGTGTGGAAGAAATGTTTGAAGCAAAAGGTTAAAGAGTGA 20959
Qy 961 AAAAAATTAATTAACAGTGTAGCTGTGTGGAAGATGGATAGGAGCTTAATAGAGAA 1020
Db 20960 AAAAAATTAATTAACAGTGTAGCTGTGTGGAAGATGGATAGGAGCTTAATAGAGAA 21019
Qy 1021 CAAATAGGAATTTTCAATGTAATGTAAGAACCTGTGTAAGGTTGAAGCAATTA 1080
Db 21020 CAAATAGGAATTTTCAATGTAATGTAAGAACCTGTGTAAGGTTGAAGCAATTA 21079
Qy 1081 TCTGCAACATCGCGGACTTTTTTTTTTATGTAATGCTTGGCAATTTAAATAGAGAGA 1140
Db 21080 TCTGCAACATCGCGGACTTTTTTTTTTATGTAATGCTTGGCAATTTAAATAGAGAGA 21139
Qy 1141 GAGAAATGTAAGAGTGTGATAGTCAAGAGTGAAGTCTGCCAGACATGTGAAGAG 1200
Db 21140 GAGAAATGTAAGAGTGTGATAGTCAAGAGTGAAGTCTGCCAGACATGTGAAGAG 21199
Qy 1201 AGACAGTGAAGAGAGATTTATCAAGACAGAAAGTAAATGCTGCGCAGT 1260
Db 21200 AGACAGTGAAGAGAGATTTATCAAGACAGAAAGTAAATGCTGCGCAGT 21259

Qy 1261 GGCATCTAGTCTGAGTCTTAATCTGAGGGAAGAGATGAAGATTAAGAGCTTGTCTAGT 1320
Db 21260 GGCATCTAGTCTGAGTCTTAATCTGAGGGAAGAGATGAAGATTAAGAGCTTGTCTAGT 21319
Qy 1321 TATGAAGAGAGTGAAGGCTTCAAGAGCTTACAGGCTTGAATTAATAGAAATGATTG 1380
Db 21320 TATGAAGAGAGTGAAGGCTTCAAGAGCTTACAGGCTTGAATTAATAGAAATGATTG 21379
Qy 1381 GAGAAAGATTAATCTGAGAGAGTGAATTTCAAGCTTGAATGATCTCAATACAGA 1440
Db 21380 GAGAAAGATTAATCTGAGAGAGTGAATTTCAAGCTTGAATGATCTCAATACAGA 21439
Qy 1441 CACTGTCTAAATGCTTCAAGACATGATCCCTGCCCTCAAGGACTTACAGCAAAAC 1500
Db 21440 CACTGTCTAAATGCTTCAAGACATGATCCCTGCCCTCAAGGACTTACAGCAAAAC 21499
Qy 1501 AAGAGTAAAGAAATCACCCAACTAATTAAGCACTGTGTGAATATCAAGAAAG 1560
Db 21500 AAGAGTAAAGAAATCACCCAACTAATTAAGCACTGTGTGAATATCAAGAAAG 21559
Qy 1561 AAATAGATCTAGTACTGTAGATGTGCAACGCGCATCAAGATATCTTCTAGTTTCAAGA 1620
Db 21560 AAATAGATCTAGTACTGTAGATGTGCAACGCGCATCAAGATATCTTCTAGTTTCAAGA 21619
Qy 1621 GTTTCAGATCGCGCGGCGGCTGCTCAGCGCTGTATCCACAGCATTTGGAGGCGCA 1680
Db 21620 GTTTCAGATCGCGCGGCGGCTGCTCAGCGCTGTATCCACAGCATTTGGAGGCGCA 21679
Qy 1681 GCGGCGGTGATCAAGAGTCAAGAGATCAAGACATCTCGTGAACAGCGTGAACCCCG 1740
Db 21680 GCGGCGGTGATCAAGAGTCAAGAGATCAAGACATCTCGTGAACAGCGTGAACCCCG 21739
Qy 1741 TCTCTACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1800
Db 21740 TCTCTACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 21799
Qy 1801 ACTCAGAGGCTGAGGCAAGAGATGCGGTGAACCCGGAGGTGAATGTTGCTGAGCG 1860
Db 21800 ACTCAGAGGCTGAGGCAAGAGATGCGGTGAACCCGGAGGTGAATGTTGCTGAGCG 21859
Qy 1861 AGATGCGGCACTGGGCTCCAGCTGGGGAAGAGTGAATGCTGCTCAAAAAA 1920
Db 21860 AGATGCGGCACTGGGCTCCAGCTGGGGAAGAGTGAATGCTGCTCAAAAAA 21919
Qy 1921 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1980
Db 21920 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 21979
Qy 1981 AGCATGTTACAGGAGAGAAAAATGTTTTCAGCAAGATGACAGATGAGGAAATAGAGAT 2040
Db 21980 AGCATGTTACAGGAGAGAAAAATGTTTTCAGCAAGATGACAGATGAGGAAATAGAGAT 22039
Qy 2041 ATGTTCAAGGAAGAGACCCAGATCATGCTTTTGAAGGTTAAGGAACAAGTGT 2100
Db 22040 ATGTTCAAGGAAGAGACCCAGATCATGCTTTTGAAGGTTAAGGAACAAGTGT 22099
Qy 2101 TTGCAATCTCAGGTTCAATTAATGATGATTAATGATGATGATGATGATGATGATGAT 2160
Db 22100 TTGCAATCTCAGGTTCAATTAATGATGATTAATGATGATGATGATGATGATGATGAT 22159
Qy 2161 TTTTAAAAATGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 22160 TTTTAAAAATGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22219
Qy 2221 TTGAAGAGCAAGTATTTTCTGCACTTTTGTCCAACTTTGATGATGATGATGATGATGAT 2280
Db 22220 TTGAAGAGCAAGTATTTTCTGCACTTTTGTCCAACTTTGATGATGATGATGATGATGAT 22279
Qy 2281 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 22280 TTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22339

QY	2241	AATCTTTAAAGCACTGGCCCTGGTCTAACTTCTTATTTTGCAGAGGAATTCGAATC	2400
Db	22340	AATCTTTAAAGCACTGGCCCTGGTCTAACTTCTTATTTTGCAGAGGAATTCGAATC	22399
QY	2401	TGAGAGCAAAACATTTTGGCTCGAGTTAATGAAACAGCTTATGGCTTCTTAAAGTGA	2460
Db	22400	TGAGAGCAAAACATTTTGGCTCGAGTTAATGAAACAGCTTATGGCTTCTTAAAGTGA	22455
QY	2461	TTCTTTAGTTAAATTTCTTTCCCACTAGTGCATATCTGCACTTCTAGTTCTGTGGCTGA	2520
Db	22460	TTCTTTAGTTAAATTTCTTTCCCACTAGTGCATATCTGCACTTCTAGTTCTGTGGCTGA	22515
QY	2521	AATACAGAAATATTTAGTAAACAGCATACAAAGTCTGGGGAAATATATTTGGGTAGTG	2580
Db	22520	AATACAGAAATATTTAGTAAACAGCATACAAAGTCTGGGGAAATATATTTGGGTAGTG	22577
QY	2581	GCTGAGAGCCCTCATTTTCTTAAGAATGTGACCTTTAGCAGGGTATGTGGCTCACCT	2640
Db	22580	GCTGAGAGCCCTCATTTTCTTAAGAATGTGACCTTTAGCAGGGTATGTGGCTCACCT	22633
QY	2641	ATAATTCAGACATTTTGGAGGCCAATCTCAAGAAATCGCTTGAAACCAAGATCTCAGA	2700
Db	22640	ATAATTCAGACATTTTGGAGGCCAATCTCAAGAAATCGCTTGAAACCAAGATCTCAGA	22699
QY	2701	CTAGCATGGGCMAATATGCAAGACCTCATCTCTACAAAAATTTAAAAATCAGCTGAGCA	2760
Db	22700	CTAGCATGGGCMAATATGCAAGACCTCATCTCTACAAAAATTTAAAAATCAGCTGAGCA	22755
QY	2761	TGTGTGCAATAGCCCTGTATGTCCCACTTACCTGGGAACTAGTGGGTGATCCGCTTGACA	2820
Db	22760	TGTGTGCAATAGCCCTGTATGTCCCACTTACCTGGGAACTAGTGGGTGATCCGCTTGACA	22815
QY	2821	CAGAGATTTGAGGCTTAAAGTGTGAGCGCATGATCACAACTGCATCTCAAGCTTGTAGTGACAG	2880
Db	22820	CAGAGATTTGAGGCTTAAAGTGTGAGCGCATGATCACAACTGCATCTCAAGCTTGTAGTGACAG	22877
QY	2881	AGGAAGACCTGTCCCTTAAAAAGAAAGAAATGTGATTTTATTCCTTTAGCAGTACAGT	2940
Db	22880	AGGAAGACCTGTCCCTTAAAAAGAAAGAAATGTGATTTTATTCCTTTAGCAGTACAGT	22933
QY	2941	CATTAGTCATTAAGTTTGGAGTTTGAGAGAAATTAATATGATCAGAAAGAAATTTATATCACT	3000
Db	22940	CATTAGTCATTAAGTTTGGAGTTTGAGAGAAATTAATATGATCAGAAAGAAATTTATATCACT	22999
QY	3001	GTGTCTGTAGAGATATATGAAGAAATTAAGAGACTAGAGTCAGGAGATTCACCTTAAGTG	3060
Db	23000	GTGTCTGTAGAGATATATGAAGAAATTAAGAGACTAGAGTCAGGAGATTCACCTTAAGTG	23055
QY	3061	TTTGTGTCTTTGTTTGAAGAGAGTCTTTTGTATCCAGAGCTAGAGTGCAATGGTG	3120
Db	23060	TTTGTGTGTGTGTGTGAAGAGAGTCTTTTGTATCCAGAGCTAGAGTGCAATGGTG	23115
QY	3121	CAGTCATGGCTCACCGCAGCTCAAACTCCGAGCTCAAAATTAATCTTCCGAGCTGGCCT	3180
Db	23120	CAGTCATGGCTCACCGCAGCTCAAACTCCGAGCTCAAAATTAATCTTCCGAGCTGGCCT	23177
QY	3181	CCCAAAAGTCTGGAATTAAGGTGTAGGCCAAAGGTTTATGTATGTGTCTGGCTTAAGT	3240
Db	23180	CCCAAAAGTCTGGAATTAAGGTGTAGGCCAAAGGTTTATGTATGTGTCTGGCTTAAGT	23233
QY	3241	GCCTCTCAAACTTCAAGTACGAGACAAGTGCAGGAACTCGGAACTCGACTCAACAAAGTCTGGGTT	3300
Db	23240	GCCTCTCAAACTTCAAGTACGAGACAAGTGCAGGAACTCGGAACTCGACTCAACAAAGTCTGGGTT	23299
QY	3301	TAAAGCTGAGCCTTGCAATTTCAATGATGCTCAAGCTGATGCAAGTGTGTCTGATCAAGA	3360
Db	23300	TAAAGCTGAGCCTTGCAATTTCAATGATGCTCAAGCTGATGCAAGTGTGTCTGATCAAGA	23355
QY	3361	GCCAAAGCACTGAGCAGCAAGAGATCTAGTTTACATTTGTATTCGAAGTTGATATTAATAGT	3420
Db	23360	GCCAAAGCACTGAGCAGCAAGAGATCTAGTTTACATTTGTATTCGAAGTTGATATTAATAGT	23415
QY	3421	AGTGCAATTAAGATGAGAGGAATGTGMAAATCAGTAACMAAGAAAGTTTCACTCTTG	3480

Db	23420	AGTGCATTAAGATGAGAGGAAATGTGAAATACGTAACTAAAGAAAGTTCAACCTCTTG	23479
Qy	3481	GTAAATGTAGCATGAGAGGAAAGGATGGGGCAAACTAACTGTGTTTGTGTGTACT	3540
Db	23480	GTAAATGTAGCATGAGAGGAAAGGATGGGGCAAACTAACTGTGTTTGTGTGTACT	23539
Qy	3541	GAGAGGAGAAATTGTAGCTCTAATTAACAGAAATAGGAGAAAGATGGTTTGAGAGAA	3600
Db	23540	GACAGAGAAATTTGTAGCTCTAATTAACAGAAATAGGAGAAAGATGGTTTGAGAGAA	23599
Qy	3601	GAGAGTCTGTTCAGACGTGTGAGGTCCTCAGGTGAGACAGGATCTCCAAAGGGAAT	3660
Db	23600	GAGGAGTCTGTTCAGACGTGTGAGGTCCTCAGGTGAGACAGGATCTCCAAAGGGAAT	23658
Qy	3661	GAGCATAGGCAACTTAAAGAAATCTGTGCTCAAGGGAGCTGTGAGCTCGAGTGT	3720
Db	23660	GAGCATAGGCAACTTAAAGAAATCTGTGCTCAAGGGAGCTGTGAGCTCGAGTGT	23719
Qy	3721	AGATCTGAGGGTCAATGACACATAGAGTTTGAAGAACAAGAGTAGGCAACAAAGAC	3780
Db	23720	AGATCTGAGGGTCAATGACACATAGAGTTTGAAGAACAAGAGTAGGCAACAAAGAC	23779
Qy	3781	AAATACAAAGAGAGGAGACTGATGATGAGACTTTTGCTTTTNGATGAGAGAGGA	3840
Db	23780	AAATACAAAGAGAGGAGACTGATGATGAGACTTTTGCTTTTNGATGAGAGAGGA	23839
Qy	3841	ACAGGAAATAGGAATGAAGGGAACAGCTTGTAGAAATGTAGACATCTGAAAAA	3900
Db	23840	ACAGGAAATAGGAATGAAGGGAACAGCTTGTAGAAATGTAGACATCTGAAAAA	23899
Qy	3901	ATACACATGTCTATGGAAGTCAAGGGAAGAGATTTCAAGAGAGGATATGTGTGACA	3960
Db	23900	ATACACATGTCTATGGAAGTCAAGGGAAGAGATTTCAAGAGAGGATATGTGTGACA	23959
Qy	3961	GTATTAACAACATCAGGAATATACAGCTTAAAGTATATCTGTGACATGACCTGTGG	4020
Db	23960	GTATTAACAACATCAGGAATATACAGCTTAAAGTATATCTGTGACATGACCTGTGG	24015
Qy	4021	ATTGTGAGGAGACACTAATAATAAGGAATTTATGTGGTATATGAGGCAAAAG	4080
Db	24020	ATTGTGAGGAGACACTAATAATAAGGAATTTATGTGGTATATGAGGCAAAAG	24078
Qy	4081	GAAAGGTTATCCAAATCAAAAGCAAGGTGGGAGTAAAGGATGATTTTCCAAAGTGAGGA	4140
Db	24080	GAAAGGTTATCCAAATCAAAAGCAAGGTGGGAGTAAAGGATGATTTTCCAAAGTGAGGA	24133
Qy	4141	TCAGTGAATGGGGAAGGAGCAGAGACATCCATGGCCATCCAGGCAAGCCACCTTCA	4200
Db	24140	TCAGTGAATGGGGAAGGAGCAGAGACATCCATGGCCATCCAGGCAAGCCACCTTCA	24199
Qy	4201	GAACTTCATGAGAGTTCAAGCTATCCAGAAAGTCTCTGATACCTTATCTTTCTGGGTT	4260
Db	24200	GAACTTCATGAGAGTTCAAGCTATCCAGAAAGTCTCTGATACCTTATCTTTCTGGGTT	24255
Qy	4261	TGCAATAGGCTCATTTGTATGGGATGATTTAATTAATATATGGCCACTGGTATCAACT	4320
Db	24260	TGCAATAGGCTCATTTGTATGGGATGATTTAATTAATATATGGCCACTGGTATCAACT	24313
Qy	4321	AACCTTCAACCCCTCTCCCTCCCTATCATGCTTGCTTTTCCAGTACCACTGCCCTA	4380
Db	24320	AACCTTCAACCCCTCTCCCTCCCTATCATGCTTGCTTTTCCAGTACCACTGCCCTA	24378
Qy	4381	TCCATAGCTACCCCATAGGTGTGACAGTATCACTCACTCAAAAAGCATCACTTGG	4440
Db	24380	TCCATAGCTACCCCATAGGTGTGACAGTATCACTCACTCAAAAAGCATCACTTGG	24433
Qy	4441	GAGATTTCTAAGAAATTTAGAGTGTGCTGCAGAGAAATTTAGTTGAAGATCAATATATAT	4500
Db	24440	GAGATTTCTAAGAAATTTAGAGTGTGCTGCAGAGAAATTTAGTTGAAGATCAATATATAT	24499
Qy	4501	TTCAATATATCACTGTGTCTATTTTATATCAGGGCCATTAATATGTTTTAAACAAG	4560

Db 24500 TTCACAAATACAGTCCGTGCTATTTTATATACAGGCCCATTAATAATGTTTAAACAAAG 24559
 QY 4561 AGGTGATAAATTCAGATTTTCTTTTATTAAGCTTACACGTGATGACAGTGGTGAATAG 4620
 Db 24560 AGGTGATAAATTCAGATTTTCTTTTATTAAGCTTACACGTGATGACAGTGGTGAATAG 24619
 QY 4621 ATTGGGATGAGGGCAATCTTTTATTAAGCTTATATTCCTCCGACCCCTACTTCTC 4680
 Db 24620 ATTGGGATGAGGGCAATCTTTTATTAAGCTTATATTCCTCCGACCCCTACTTCTC 24679
 QY 4681 CTGTTTCTTCTACCTCTCTCCCTACTACACAGAAACCTTCTCTCCCTACTCAT 4740
 Db 24680 CTGTTTCTTCTACCTCTCTCCCTACTACACAGAAACCTTCTCTCCCTACTCAT 24739
 QY 4741 TCCCTGAATGCTGCTGCTGTTAGAGTTCCAGCCTTGAACGTGAGGCTAATCAGAACAC 4800
 Db 24740 TCCCTGAATGCTGCTGCTGTTAGAGTTCCAGCCTTGAACGTGAGGCTAATCAGAACAC 24739
 QY 4801 AGTGTACAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4860
 Db 24800 AGTGTACAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24859
 QY 4861 GCCGGAACGACGGGAAAAAACAAGATGACGATATTTTCTGATCTGATTTAA 4920
 Db 24860 GCCGGAACGACGGGAAAAAACAAGATGACGATATTTTCTGATCTGATTTAA 24919
 QY 4921 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4980
 Db 24920 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24979
 QY 4981 ATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5040
 Db 24980 ATACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25039
 QY 5041 GCTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5100
 Db 25040 GCTCAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25099
 QY 5101 AATCATCAGAAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5160
 Db 25100 AATCATCAGAAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25159
 QY 5161 ATGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5220
 Db 25160 ATGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 25219
 QY 5221 TTTATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 5280
 Db 25220 TTTATATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 25279
 QY 5281 ACGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5340
 Db 25280 ACGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25339
 QY 5341 ACCTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5400
 Db 25340 ACCTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25399
 QY 5401 CAGGACCCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5460
 Db 25400 CAGGACCCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25459
 QY 5461 ATGTGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5520
 Db 25460 ATGTGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25519
 QY 5521 AAAGTGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5580
 Db 25520 AAAGTGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25579
 QY 5581 ATGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5640
 Db 25580 ATGGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25639

QY 5641 GCCTTGCCCTCTCAAGATTTCTGGGATTACATGATGATCCCTGCGCCCTGCGCCAGAAATAT 5700
 Db 25640 GCCTTGCCCTCTCAAGATTTCTGGGATTACATGATGATCCCTGCGCCCTGCGCCAGAAATAT 25699
 QY 5701 CAATATATTTAGATTTTATGAGAACTGATGAGAACTGATGAGAACTGATGAGAACTGATGAGAA 5760
 Db 25700 CAATATATTTAGATTTTATGAGAACTGATGAGAACTGATGAGAACTGATGAGAACTGATGAGAA 25759
 QY 5761 AAAAATTAGAGATTTGGAACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5820
 Db 25760 AAAAATTAGAGATTTGGAACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25819
 QY 5821 ATGATGCCAATGCTGCAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 5880
 Db 25820 ATGATGCCAATGCTGCAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 25879
 QY 5881 GTCATGGGAAACCAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 5940
 Db 25880 GTCATGGGAAACCAAGATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 25939
 QY 5941 TATAGGGGATGATTTTGTATTTTCAAGAGACAAACAGTTCCATCAATGCGCAAGATCTTA 6000
 Db 25940 TATAGGGGATGATTTTGTATTTTCAAGAGACAAACAGTTCCATCAATGCGCAAGATCTTA 25999
 QY 6001 GTGTGTACACGAGGATTAAGTGTCTGAAAGTGAATTTGGAAGACAGATCATTTGAGCTGA 6060
 Db 26000 GTGTGTACACGAGGATTAAGTGTCTGAAAGTGAATTTGGAAGACAGATCATTTGAGCTGA 26059
 QY 6061 GGTGGCTAGAGCTGCTTCATGAGACATTAAGTGAATTTGGAAGACAGATCATTTGAGCTGA 6120
 Db 26060 GGTGGCTAGAGCTGCTTCATGAGACATTAAGTGAATTTGGAAGACAGATCATTTGAGCTGA 26119
 QY 6121 TGCCCACTCTTCAGTGAATGACAG 6180
 Db 26120 TGCCCACTCTTCAGTGAATGACAG 26179
 QY 6181 GAAAGATTAAGAGATGCTG 6201
 Db 26180 GAAAGATTAAGAGATGCTG 26200

RESULT 3
 AC007242/c 112748 bp DNA linear PRI 30-SEP-2000
 LOCUS
 DEFINITION Homo sapiens BAC clone RP11-182H9 from 2, complete sequence.
 AC007242
 VERSION AC007242.3 GI:5931465
 KEYWORDS
 HTG.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 112748)
 Sulston, J.E. and Waterston, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 JOURNAL MEDLINE 99063792
 PUBMED 9847074
 REFERENCE
 2 (bases 1 to 112748)
 Kozlowicz, A., Stoneking, T., Hawkins, M. and Hawrylyko, C.
 The sequence of Homo sapiens BAC clone RP11-182H9
 JOURNAL Unpublished
 REFERENCE
 3 (bases 1 to 112748)
 Waterston, R.H.
 Direct Submission
 Submitted (07-APR-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 112748)
 Waterston, R.H.
 Direct Submission
 Submitted (28-SEP-1999) Genome Sequencing Center, Washington

REFERENCE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 112748)
AUTHORS Watson,R.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 28, 1999 this sequence version replaced gi:5001521.

COMMENT

Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0182H09

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,B.,
Tateno,M., Cataneese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(<http://bacpac.med.buffalo.edu>)
VECTOR: pBAC3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-53H10, 200 bp overlap, the
clone sequenced to the right is RP11-309N8. Actual start of this
clone is at base position 100126 of RP11-53H10; actual end is at
112748 of RP11-182H9.

The reads used to determine the sequence at the region of base
positions 31840 to 112748 consists of numerous polymorphic base
changes represented from both NH0182H09 and RP11-309N8. It is
unknown which base calls belong to which clone.

FEATURES

source

```
1. 112748
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="2"
   /map="2"
   /clone="RP11-182H9"
   /clone_1lb="RPCI-11"
   1033..1331
   /rpt_family="Alu"
   3574..3737
   /rpt_family="MIR"
   3814..4009
   /rpt_family="MER1_type"
   4048..4358
   /rpt_family="Alu"
```

```
repeat_region 6038..6457
/rpt_family="MIR"
repeat_region 6477..6680
/rpt_family="L2"
repeat_region 6681..6978
/rpt_family="Alu"
repeat_region 7386..7542
/rpt_family="Alu"
repeat_region 7543..7567
/rpt_family="(TTTTG)n"
repeat_region 7568..7702
/rpt_family="Alu"
repeat_region 8575..8596
/rpt_family="AT-rich"
repeat_region 8599..9529
/rpt_family="L1"
repeat_region 9685..9875
/rpt_family="MIR"
repeat_region 10125..10238
/rpt_family="MIR"
repeat_region 10357..10458
/rpt_family="MIR"
repeat_region 10484..10784
/rpt_family="MER1_type"
repeat_region 10898..11020
/rpt_family="Alu"
repeat_region 11835..12166
/rpt_family="MIR"
STS
/db_xref="GI:1234078"
repeat_region 12082..12118
/rpt_family="(CA)n"
repeat_region 12210..12521
/rpt_family="Alu"
repeat_region 13068..13394
/rpt_family="L2"
repeat_region 13429..13479
/rpt_family="Alu"
repeat_region 13511..13812
/rpt_family="Alu"
repeat_region 13967..14092
/rpt_family="L2"
repeat_region 14111..14139
/rpt_family="AT-rich"
repeat_region 14206..14515
/rpt_family="Alu"
repeat_region 14516..14559
/rpt_family="Alu"
repeat_region 14723..14767
/rpt_family="AT-rich"
repeat_region 15360..15575
/rpt_family="MIR"
repeat_region 15691..15734
/rpt_family="L2"
repeat_region 16277..16577
/rpt_family="Alu"
repeat_region 16608..16649
/rpt_family="L2"
repeat_region 16793..16904
/rpt_family="MIR"
repeat_region 16919..17003
/rpt_family="MIR"
repeat_region 19659..19795
/rpt_family="GA-rich"
repeat_region 20648..20707
/rpt_family="L2"
repeat_region 20881..21187
/rpt_family="Alu"
repeat_region 22002..22304
/rpt_family="Alu"
repeat_region 23542..23659
/rpt_family="L2"
repeat_region 23944..24249
/rpt_family="Alu"
repeat_region 25966..26170
```

repeat_region /rpt_family="L1"
26177..26461
repeat_region /rpt_family="Alu"
26476..26808
repeat_region /rpt_family="MALR"
26809..27119
repeat_region /rpt_family="Alu"
27120..27165
repeat_region /rpt_family="MALR"
27186..27530
repeat_region /rpt_family="L1"
27858..27971
repeat_region /rpt_family="Alu"
28007..28310
repeat_region /rpt_family="Alu"
28872..29076
repeat_region /rpt_family="MIR"
29419..29758
repeat_region /rpt_family="L1"
30018..30285
repeat_region /rpt_family="Alu"
30648..30739
misc_feature /note="match to EST AA436054 (NID:92140968) zu01c12.r1"
31066..31159
repeat_region /rpt_family="MER1_type"

Query Match 98.8%; Score 6129.2; DB 9; Length 112748;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 6176; Conservative 0; Mismatches 23; Indels 3; Gaps 3;

QY 1 CACCATGGGCTCAGGCAATTTATGAAAGCCAAATATACAGCCTTAATAATGATGAGAC 60
DB 53580 CACCATGGGCTCAGGCAATTTATGAAAGCCAAATATACAGCCTTAATAATGATGAGAC 53521
QY 61 CTAAATACCAGAGAAGCTCCCTTGTGAAGATTGTGAACAAAATTAAATGATGAGAG 120
DB 53520 CTAAATACCAGAGAAGCTCCCTTGTGAAGATTGTGAACAAAATTAAATGATGAGAG 53461
QY 121 TTAAATAGTTCTAATGAGATGTGAACCCAAAGCCATATCAGGCTAGCAAAATGGCAGA 180
DB 53460 TTAAATAGTTCTAATGAGATGTGAACCCAAAGCCATATCAGGCTAGCAAAATGGCAGA 53401
QY 181 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTAGCGCTTAATGATGCTAAAGAAA 240
DB 53400 ATTCAATATCATCAAAAGTTATCTTCAAGAGCTTAGCGCTTAATGATGCTAAAGAAA 53341
QY 241 ATGTGAAGAGCCCTCAGCATCGAAGAGCAGGTGTACAGCAATGATCAAAAAGAAAA 300
DB 53340 ATGTGAAGAGCCCTCAGCATCGAAGAGCAGGTGTACAGCAATGATCAAAAAGAAAA 53281
QY 301 CCACAGGCGCTTCCCTTCCCATCTGATGATGAAGCAGTCTTCAATTTCCATAGTAGT 360
DB 53280 CCACAGGCGCTTCCCTTCCCATCTGATGATGAAGCAGTCTTCAATTTCCATAGTAGT 53221
QY 361 AAATTTTCTAGATACAGCTTGTAGAGCTTAAGAGCTGAAGAAAGAAAGCTCCATCAAG 420
DB 53220 AAATTTTCTAGATACAGCTTGTAGAGCTTAAGAGCTGAAGAAAGAAAGCTCCATCAAG 53161
QY 421 GAAATTTATCTAAGTACTGTAAATGATCAATTTTGTGATTTGATTTGAATTAATTAAGT 480
DB 53160 GAAATTTATCTAAGTACTGTAAATGATCAATTTTGTGATTTGATTTGAATTAATTAAGT 53101
QY 481 TGTAGCCTGCGCGGTGCTCAGCGCTGTAATCCAGAGCTTTGGAGGCGCAGAGTGGG 540
DB 53100 TGTAGCCTGCGCGGTGCTCAGCGCTGTAATCCAGAGCTTTGGAGGCGCAGAGTGGG 53041
QY 541 CAGATCATAGAGTCAGAGATTGAGACAGAGCTTACAGCAATGATGAAAAGCTCCGCTCTTA 600
DB 53040 CAGATCATAGAGTCAGAGATTGAGACAGAGCTTACAGCAATGATGAAAAGCTCCGCTCTTA 52981
QY 601 CTAAAGATCAAAAATTAGCAGAGTGTGTGGCGCAGACCTGTAAAGCCAGCTGTCTGA 660
DB 52980 CTAAAGATCAAAAATTAGCAGAGTGTGTGGCGCAGACCTGTAAAGCCAGCTGTCTGA 52921

QY 661 GAGAGTGAAGCAGAGAAATGCTTGAACCCAGAGGCAAGGTGCAAGCAAGATCA 720
DB 52920 GAGAGTGAAGCAGAGAAATGCTTGAACCCAGAGGCAAGGTGCAAGCAAGATCA 52861
QY 721 CACCAATGCACTGTAGCCTGATGACAGGGCAAGACTCCCACTCAAAAAAAGAAAAA 780
DB 52860 CACCAATGCACTGTAGCCTGATGACAGGGCAAGACTCCCACTCAAAAAAAGAAAAA 52801
QY 781 AAGAAATATGTAAGTTGTGCTATTAACAAATTAATAGCAGTGAAGCAAGTGTAAAG 840
DB 52800 AAGAAATATGTAAGTTGTGCTATTAACAAATTAATAGCAGTGAAGCAAGTGTAAAG 52741
QY 841 CCTATGACCATGTGTAATCTGAGAACTAGTGGAAACACATATTAAGGAACTTAAGCAAT 900
DB 52740 CCTATGACCATGTGTAATCTGAGAACTAGTGGAAACACATATTAAGGAACTTAAGCAAT 52681
QY 901 CCTGGAAGTAAAGTTTGGAAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTGAAGAA 960
DB 52680 CCTGGAAGTAAAGTTTGGAAAGAAATGTTTGAAGCAAAAGGTTTAAAGAGTGAAGAA 52621
QY 961 AAAAAATTAATACAGTTTACCTGTGTGAGAAATGGATAGGAGCTAATAGAGAAAT 1020
DB 52620 AAAAAATTAATACAGTTTACCTGTGTGAGAAATGGATAGGAGCTAATAGAGAAAT 52561
QY 1021 CAATATGAAATGTTTCATGTAATGTTAAGAGACCTGTGTAAGGCTGAAGCAATTAATTA 1080
DB 52560 CAATATGAAATGTTTCATGTAATGTTAAGAGACCTGTGTAAGGCTGAAGCAATTAATTA 52501
QY 1081 TCTGACCATGCGGGGACTTTTGTATGTAAGTGTGCAATTTAAATAGAGAGCA 1140
DB 52500 TCTGACCATGCGGGGACTTTTGTATGTAAGTGTGCAATTTAAATAGAGAGCA 52441
QY 1141 GAGAAATGTAAGTGTGATGATGAGTGAAGTTCTGACAGCATGTGAAGAAAG 1200
DB 52440 GAGAAATGTAAGTGTGATGATGAGTGAAGTTCTGACAGCATGTGAAGAAAG 52381
QY 1201 AGACAGTGAAGCAAGAGTGAAGATTAACAAGACAGATTAATGTCTGCGCAAT 1260
DB 52380 AGACAGTGAAGCAAGAGTGAAGATTAACAAGACAGATTAATGTCTGCGCAAT 52321
QY 1261 GGCATCTAGTCTGAGCTTAATCTGAGGAAGAAAGTGAAGATTAAGTCTGATAT 1320
DB 52320 GGCATCTAGTCTGAGCTTAATCTGAGGAAGAAAGTGAAGATTAAGTCTGATAT 52261
QY 1321 TATGAAGAGATGGAAGCTTCAAGACCTACAGGTTGATTAATAGAAATGATG 1380
DB 52260 TATGAAGAGATGGAAGCTTCAAGACCTACAGGTTGATTAATAGAAATGATG 52201
QY 1381 GAGAAAGAAATTAATCTGAGAGAGTGAATTTCAAGCTTGAAGTCTACATACAGA 1440
DB 52200 GAGAAAGAAATTAATCTGAGAGAGTGAATTTCAAGCTTGAAGTCTACATACAGA 52141
QY 1441 CACTGTGCTAATGCTTCAAAAGCATGATCCCTGCTCAAGGACTTACACCAAAAAC 1500
DB 52140 CACTGTGCTAATGCTTCAAAAGCATGATCCCTGCTCAAGGACTTACACCAAAAAC 52081
QY 1501 AAGAGTAAAGAAATCAACCAATCTATTATAGACACTTGTAGAAATATCAAGAAAG 1560
DB 52080 AAGAGTAAAGAAATCAACCAATCTATTATAGACACTTGTAGAAATATCAAGAAAG 52021
QY 1561 AAATACGATCTAGTACTGATAGTGAACGGCATCAAGATATCTTCAATTTCAAGAA 1620
DB 52020 AAATACGATCTAGTACTGATAGTGAACGGCATCAAGATATCTTCAATTTCAAGAA 51961
QY 1621 GTTTCAGATCGGCGGGGCGGTGCTCAGCGCTGTATCCAGACATTTGGAGGCCGA 1680
DB 51960 GTTTCAGATCGGCGGGGCGGTGCTCAGCGCTGTATCCAGACATTTGGAGGCCGA 51901
QY 1681 GCGGGGTGATCACAAGGTCAGAGATCAAGACCATCTGCTTAAACAAGTGAAGCCCG 1740
DB 51900 GCGGGGTGATCACAAGGTCAGAGATCAAGACCATCTGCTTAAACAAGTGAAGCCCG 51841

QY	1741	TTCTCTACAAAAAATATATAAAAAATATAGCCAGCGGTGGGGCGCTGTGAATGCCAGCT	1800
Db	51840	TTCTTACTTAAAAATATATAAAAAATTAGCCAGCGGTGGGGCGCTGTGAATGCCAGCT	51789
QY	1801	ACTCAGGAGGCTGAGCGAGGAGATGGCTGTGAACCCCGGAGGTAGATTGTCCGTGACCG	1860
Db	51780	ACTCAGGAGGCTGAGCGAGGAGATGGCTGTGAACCCCGGAGGTAGATTGTCCGTGACCG	51722
QY	1861	AGATCGGCGCACCTGCCGCTCCAGCCTGGGGCGACAGATGAGACCTGCCCTTCAAAAAAAA	1920
Db	51720	AGATCGGCGCACCTGCCGCTCCAGCCTGGGGCGACAGATGAGACCTGCCCTTCAAAAAAAA	51665
QY	1921	AAAAAAGTTTCAAGTCTTAAACAACCTGATTTCAACAGCTCTAGAAATAGAG	1980
Db	51661	AAAAAAGTTTCAAGTCTTAAACAACCTGATTTCAACAGCTCTAGAAATAGAG	51602
QY	1981	AGCATGTTTACAGGAGAGAAAATGTTTTTCAGCAAGGTACAGATAGGGAATAGAGAT	2040
Db	51601	AGCATGTTTACAGGAGAGAAAATGTTTTTCAGCAAGGTACAGATAGGGAATAGAGAT	51544
QY	2041	ATGTTTCAAGGAAGAGAGACCCAGAGTCATGTTTTGTAGGGTTAGAGGAACAAGTGT	2100
Db	51541	ATGTTTCAAGGAAGAGAGACCCAGAGTCATGTTTTGTAGGGTTAGAGGAACAAGTGT	51488
QY	2101	TTGCAATCTCCAGGTTCCATTATGTCGTTATGAAATCAATATGATGTAGCAACTGCA	2160
Db	51481	TTGCAATCTCCAGGTTCCATTATGTCGTTATGAAATCAATATGATGTAGCAACTGCA	51422
QY	2161	TTTTTAAAAAATGAAATTAATAGAGAGAAAGAAATGAAAAATTTAGCAGCTTTCATT	2220
Db	51421	TTTTTAAAAAATGAAATTAATAGAGAGAAAGAAATGAAAAATTTAGCAGCTTTCATT	51365
QY	2221	TTGAAAGAGCAGATTAATATTTCTGCAACTTTTGCTTCAATTGTAACGTACTTAATATT	2280
Db	51361	TTGAAAGAGCAGATTAATATTTCTGCAACTTTTGCTTCAATTGTAACGTACTTAATATT	51302
QY	2281	TTATGTATGATGTGAAATTAACAGTATACATATATTTCTTACTGTAGACTGCACTAAAA	2340
Db	51301	TTATGTATGATGTGAAATTAACAGTATACATATATTTCTTACTGTAGACTGCACTAAAA	51242
QY	2341	AAATCTTTAAAGCACTGGCTGTGTAACTTCTTAATTTTGACAGAGAGAAATCCAAGTC	2400
Db	51241	AAATCTTTAAAGCACTGGCTGTGTAACTTCTTAATTTTGACAGAGAGAAATCCAAGTC	51188
QY	2401	TGAGAGAGCAAACTTTTGCTGTAGTTATGAAACCACTTATGCCATTGCTAAAAAGTGA	2460
Db	51181	TGAGAGAGCAAACTTTTGCTGTAGTTATGAAACCACTTATGCCATTGCTAAAAAGTGA	51122
QY	2461	TTCTTAGTAAAAATTTCTTCCCACTAGTGCAATCTGCACTGTGAGGCTGTA	2520
Db	51121	TTCTTAGTAAAAATTTCTTCCCACTAGTGCAATCTGCACTGTGAGGCTGTA	51065
QY	2521	AAATACAGAAATATATTAAGTGAAGCAGATACAAAGTCTGGGGAATATATTTGGGTAGGT	2580
Db	51061	AAATACAGAAATATATTAAGTGAAGCAGATACAAAGTCTGGGGAATATATTTGGGTAGGT	51002
QY	2581	GCTAGAGGCTCATTTTCTTAAGAAATGTGACCTTAGGCAAGGATATGTGTGCTCACACT	2640
Db	51001	GCTAGAGGCTCATTTTCTTAAGAAATGTGACCTTAGGCAAGGATATGTGTGCTCACACT	50944
QY	2641	ATTAATTCAGACCTTTGGGAGGCCAAGTCAAGAAGATCGCTTGAAACCCAAGCTTCAGA	2700
Db	50941	ATTAATTCAGACCTTTGGGAGGCCAAGTCAAGAAGATCGCTTGAAACCCAAGCTTCAGA	50888
QY	2701	CTAGCATGGGCAATATGCAAGACCTCATCTCTACAAAAATTTAAAAATCAGCTGAGCA	2760
Db	50881	CTAGCATGGGCAATATGCAAGACCTCATCTCTACAAAAATTTAAAAATCAGCTGAGCA	50822
QY	2761	TGGTGGCACTAAGCTGTAGTCCCACTACCTGGGAGCTATGGGGGTGAGATCGCTTGA - C	2819
Db	50821	TGGTGGCACTAAGCTGTAGTCCCACTACCTGGGAGCTATGGGGGTGAGATCGCTTGAAGC	50766
QY	2820	ACAGAGATTTTAGGCTAAGGTAGACCATGATACACAACACTGCACCTCAGCTTGAAGTACA	2879

Db	50761	ACAGAGTTTGAGGCTTAAGGTCAGACCATTGATCACAACAATGCACCTCCAGCTTGAGTGACA	50702
Qy	2880	GAGGAAGACCCCTGTCCCTTAAAAAGAAATGTGGATTTTATTTCTTAAACAGTACAG	2939
Db	50701	GAGGAAGACCCCTGTCCCTTAAAAAGAAATGTGGATTTTATTTCTTAAACAGTACAG	50642
Qy	2940	TCATTAGCATTAAGTTTGAAGTATGAGAAATATATGATCAGAAAGAAATTATATAC	2999
Db	50641	TCATTAGCATTAAGTTTGAAGTATGAGAAATATATGATCAGAAAGAAATTATATAC	50582
Qy	3000	TGTGCTCTGAGATATATGAAAGAAATTAAGACTAGAGTCAAGGATTCACCTTAAGT	3059
Db	50581	TGTGCTCTGAGATATATGAAAGAAATTAAGACTAGAGTCAAGGATTCACCTTAAGT	50522
Qy	3060	GTTTGTTTGTTTGTTTGAGACAAGTCTCTTTTGTTTACCAGGCTAGAGTCAATGCT	3119
Db	50521	GTTTGTTTGTTTGTTTGAGACAAGTCTCTTTTGTTTACCAGGCTAGAGTCAATGCT	50462
Qy	3120	GCACTCATGCTCACCGAGCCTCAAACTCCAGCCTCAATTAATCTTCCAGCTCGGCC	3179
Db	50461	GCACTCATGCTCACCGAGCCTCAAACTCCAGCCTCAATTAATCTTCCAGCTCGGCC	50402
Qy	3180	TCCCAAGTGTGGAAATTAACAGGTGTGAGCCAAAGGGTTTATGTAGTGTGTCTGGCCTAG	3239
Db	50401	TCCCAAGTGTGGAAATTAACAGGTGTGAGCCAAAGGGTTTATGTAGTGTGTCTGGCCTAG	50342
Qy	3240	TGCTCTCAAACTTCAGTAGACAGACAAGTACCGGGAACTGACTCAACAAGTCTGGGT	3299
Db	50341	TGCTCTCAAACTTCAGTAGACAGACAAGTACCGGGAACTGACTCAACAAGTCTGGGT	50282
Qy	3300	TTAAGCTGAGCCTCTGCATTTTAACTAGATCAAGCTGATGCAGATGTGTCTGCTCAAG	3359
Db	50281	TTAAGCTGAGCCTCTGCATTTTAACTAGATCAAGCTGATGCAGATGTGTCTGCTCAAG	50222
Qy	3360	AGCCAGAGACTGACAGACAGAAAGTCTGTTAGCAATTGATATCAAGGTTGATATTATG	3419
Db	50221	AGCCAGAGACTGACAGACAGAAAGTCTGTTAGCAATTGATATCAAGGTTGATATTATG	50162
Qy	3420	TAGTGAACATTAAGATGAGAGGAATGTGAATACTAGTAAACAAAGAGATTCACTCTT	3479
Db	50161	TAGTGAACATTAAGATGAGAGGAATGTGAATACTAGTAAACAAAGAGATTCACTCTT	50102
Qy	3480	GCTAATGTGACATGAGAGGGAAGAATGGGGCCAACTATTAATCTGTGTGTGTGAC	3539
Db	50101	GCTAATGTGACATGAGAGGGAAGAATGGGGCCAACTATTAATCTGTGTGTGTGAC	50042
Qy	3540	TGACGAGGAGATTTGATGCTTATTTAACAGAAATAGAGAAAGTTGGTTTGGAGAGA	3599
Db	50041	TGACGAGGAGATTTGATGCTTATTTAACAGAAATAGAGAAAGTTGGTTTGGAGAGA	49982
Qy	3600	AGAGAGTCTCTGTTTCAACCTGTTGAGGTCCTCAGGTGAGACAGATCTCCAAAGGGAAA	3659
Db	49981	AGAGAGTCTCTGTTTCAACCTGTTGAGGTCCTCAGGTGAGACAGATCTCCAAAGGGAAA	49922
Qy	3660	TGACAGATGAGCAACTTAAAGAAATCTGTGCTCAGAAAGGAGCTGTGAGCTCCAGCTG	3719
Db	49921	TGACAGATGAGCAACTTAAAGAAATCTGTGCTCAGAAAGGAGCTGTGAGCTCCAGCTG	49862
Qy	3720	TAGATCTGAGGGTCATCGACATAGATTTAGAAACAAGAGTGAAGCAACMAAGAG	3779
Db	49861	TAGATCTGAGGGTCATCGACATAGATTTAGAAACAAGAGTGAAGCAACMAAGAG	49802
Qy	3780	CAATTAACAAGAGAGAGACTGATGATGAGCTTTGCTTTAGAGTGAAGAGAGG	3839
Db	49801	CAATTAACAAGAGAGAGACTGATGATGAGCTTTGCTTTAGAGTGAAGAGAGG	49742
Qy	3840	AACAGAAATTAAGAAATGAAGGAGAGCACTTGTAGAAATGTAGAGCATCTGAAAAAAA	3899
Db	49741	AACAGAAATTAAGAAATGAAGGAGAGCACTTGTAGAAATGTAGAGCATCTGAAAAAAA	49682
Qy	3900	AATACACTGTCTCATGGAATCAAGGAGAAAGATTTCAAGAAAGAGGATGTGTGAGC	3959

D	b	49681	AATACACACTGTCAATGAACTCAAGGGAAGAAGATTTCAAGAAAGGGATGTGTGAC	49622
Q	y	3960	AGATTTCACAGCATCAGGAATACAGCTAAAGTCACTCTTGACATGCAATGCACTTGTG	4019
D	b	49621	AGATTTCACAGCATCAGGAATACAGCTAAAGTCACTCTTGACATGCAATGCACTTGTG	49566
Q	y	4020	GATTGTGAGGACACACTAATAATAAGAAATTTATGTGGTATATGAGGACAAA	4079
D	b	49562	GATTGTGAGGACACACTAATAATAAGAAATTTATGTGGTATATGAGGACAAA	49503
Q	y	4080	GGAAGAGTTATCCAAATCAAAAGAGGTGGAGTAAGGATGAATTCACAAAGGTGAAGGC	4139
D	b	49502	GGAAGAGTTATCCAAATCAAAAGAGGTGGAGTAAGGATGAATTCACAAAGGTGAAGGC	49443
Q	y	4140	ATCAGTAATAGTGGGAAGGGGACACAGACATCCATGCGCCATCCAGGCAAGCCACTCC	4199
D	b	49442	ATCAGTAATAGTGGGAAGGGGACACAGACATCCATGCGCCATCCAGGCAAGCCACTCC	49388
Q	y	4200	AGAAGCTCCATGAGAGTTCAAGATATCCAGAAGGTCTCTGTACCTTAATCTTGTGGGTT	4259
D	b	49382	AGAAGCTCCATGAGAGTTCAAGATATCCAGAAGGTCTCTGTACCTTAATCTTGTGGGTT	49322
Q	y	4260	TTGATATAGGCTTCACTGTGTAGGCAATGATTTATTAACATAATGGCCACTGTGTATCACT	4319
D	b	49322	TTGATATAGGCTTCACTGTGTAGGCAATGATTTATTAACATAATGGCCACTGTGTATCACT	49263
Q	y	4320	TAACTTCAACCCCTCTCCCTCCCTCAATATGCGCTTGTGGCTTCCAGTGAACAGTCCCT	4379
D	b	49262	TAACTTCAACCCCTCTCCCTCCCTCAATATGCGCTTGTGGCTTCCAGTGAACAGTCCCT	49202
Q	y	4380	ATCCTAAGCTAACCCCAATGATCTGCACTATCACTCAACTCACAAAAAGACATCACTTT	4439
D	b	49202	ATCCTAAGCTAACCCCAATGATCTGCACTATCACTCAACTCACAAAAAGACATCACTTT	49143
Q	y	4440	GGAGATCTTAAGATTTTAAAGATTTGGCTGTCAAGAAATTTAGTTGAAGATCAATATATA	4499
D	b	49142	GGAGATCTTAAGATTTTAAAGATTTGGCTGTCAAGAAATTTAGTTGAAGATCAATATATA	49083
Q	y	4500	TTTCACAAATATCACAGTGTGCTATTTTATATCAGAGGCGCAATTAATGGTTTTAAACAA	4559
D	b	49082	TTTCACAAATATCACAGTGTGCTATTTTATATCAGAGGCGCAATTAATGGTTTTAAACAA	49022
Q	y	4560	GAGGTGATTAATTCAGATTTTCTTTTATTAAGCTTACACTGATGACAGTGTGTGAATA	4619
D	b	49022	GAGGTGATTAATTCAGATTTTCTTTTATTAAGCTTACACTGATGACAGTGTGTGAATA	48966
Q	y	4620	GATTGGATAGGGCAATCTTTTTTTTGAATGTTATATTCCTCCGACCTTACTTTCT	4679
D	b	48962	GATTGGATAGGGCAATCTTTTTTTTGAATGTTATATTCCTCCGACCTTACTTTCT	48903
Q	y	4680	CTTGTCTTTCTTCACTCTCTCTCCCTTACTCACAGAAATTTCTTCTCCCTTACTCA	4739
D	b	48902	CTTGTCTTTCTTCACTCTCTCTCCCTTACTCACAGAAATTTCTTCTCCCTTACTCA	48843
Q	y	4740	TTCCCTGAATGCTGTGTCTGTAAAGTTCCACCTTGCACGTGAGGCTAATCAGAACCA	4799
D	b	48842	TTCCCTGAATGCTGTGTGTCTGTAAAGTTCCACCTTGCACGTGAGGCTAATCAGAACCA	48788
Q	y	4800	CAGTGTACAGATGTGAGATGATGTGTGGAGAAAGTGCACAGATTAATGACCAATTACTT	4859
D	b	48782	CAGTGTGTACAGATGTGAGATGATGTGTGGAGAAAGTGCACAGATTAATGACCAATTACTT	48722
Q	y	4860	AGCCGGAACTGACGGGAAAAACAAGATCGAGATTTTTTTTCTGTGATCGATATTA	4919
D	b	48722	AGCCGGAACTGACGGGAAAAACAAGATCGAGATTTTTTTTCTGTGATCGATATTA	48663
Q	y	4920	AATGATGATGTGTGTCCATTCATCTGTGTATAGATCAGAAAGAAAAATTTATTTTGGAGA	4979
D	b	48662	AATGATGATGTGTGTGTCCATTCATCTGTGTATAGATCAGAAAGAAAAATTTATTTTGGAGA	48603
Q	y	4980	GATACCAATGAATGTGTGTTTTAGACATGCTAAGTTTGAAGTGAATTAATGGAATGACAGCG	5039
D	b	48602	GATACCAATGAATGTGTGTTTTAGACATGCTAAGTTTGAAGTGAATTAATGGAATGACAGCG	48543

OY	5040	AGCTCCGACTGTGTGGGCTTAAAGTAGAAGGCATTCGATTTGGAGATAAAGATTTG	509
Dd	48542	AGCTCCGACTGTGTGGGCTTAAAGTAGAAGGCATTCGATTTGGAGATAAAGATTTG	48483
OY	5100	AAATCATCAGAATACCGGTGTTCATTAGAACCTGTCAAGTGGGTAAAGATAGCTTAAGGAG	5159
Dd	48482	AAATCATCAGAATACCGGTGTTCATTAGAACCTGTCAAGTGGGTAAAGATAGCTTAAGGAG	48422
OY	5160	CATGTGTAGGTGATTAACAAGATCAAAGCGGAACCTTAGAATAACAATAATGTATT	5219
Dd	48422	CATGTGTAGGTGATTAACAAGATCAAAGCGGAACCTTAGAATAACAATAATGTATT	48365
OY	5220	ATTATTTATTTATTTATGTTTTATTTTAATTTTATTTTATTTATTTATTTTAA	5279
Dd	48362	ATTATTTATTTATTTATGTTTTATTTTAATTTTATTTTATTTTATTTATTTTAA	48303
OY	5280	GACGGGAGTCCGCTCGCTGCCAGGCTGGAGTGGAGTGGCGCAACTCACCTCC	5339
Dd	48302	GATGGGAGTCCGCTCGCTGCCAGGCTGGAGTGGAGTGGCGCAACTCACCTCC	48244
OY	5340	AACCTCGCTTCCCGGTTCAAGGAGCCCTCGCTCAAGCCTCTCAAGTCTGGAGCT	5399
Dd	48242	AACCTCGCTTCCCGGTTCAAGGAGCCCTCGCTCAAGCCTCTCAAGTCTGGAGCT	48185
OY	5400	ACAGGCAACCAACCACTCACCTGACTAATTTTGTATTTTATAGAGACGGGTTTCA	5459
Dd	48182	ACAGGCAACCAACCACTCACCTGACTAATTTTGTATTTTATAGAGACGGGTTTCA	48122
OY	5460	CATGTGGCCAGGCTGGTCTTGAACCTTGTGACCTTGAAGATTCACCTGCTTGCGCTTC	5519
Dd	48122	CATGTGGCCAGGCTGGTCTTGAACCTTGTGACCTTGAAGATTCACCTGCTTGCGCTTC	48065
OY	5520	CAAAGTCTGAGGATTAACAGTATAGAGCACTGCGCCGACCTATTTTGTTTTATTA	5579
Dd	48062	CAAAGTCTGAGGATTAACAGTATAGAGCACTGCGCCGACCTATTTTGTTTTATTA	48002
OY	5580	GATGGGCTCTTGTCTATGTTGCCAGGCTGGTCTCGAACCCTTGGACTCAAGCAATCTCC	5639
Dd	48002	GATGGGCTCTTGTCTATGTTGCCAGGCTGGTCTCGAACCCTTGGACTCAAGCAATCTCC	47944
OY	5640	TGCTTTGGCTCTCAAAAGTTCTGGGATTAACAATGTAGTCCCTGCGCTGGCAGAAAT	5699
Dd	47942	TGCTTTGGCTCTCAAAAGTTCTGGGATTAACAATGTAGTCCCTGCGCTGGCAGAAAT	47883
OY	5700	TCAATATATTAATTTTATAGTAGAAGTAGAACCATAGAAAAGAACCAACGAGGGGAGAA	5759
Dd	47882	TCAATATATTAATTTTATAGTAGAAGTAGAACCATAGAAAAGAACCAACGAGGGGAGAA	47822
OY	5760	GAAGAAATTAGAGATTGTGAAACAAAGAGAGAGTGCCTCAGAGAGAAAGCATGTGTC	5819
Dd	47822	GAAGAAATTAGAGATTGTGAAACAAAGAGAGAGTGCCTCAGAGAGAAAGCATGTGTC	47765
OY	5820	TATGATGCCAAATGCTGCAAAAGATAAGAAATAAGATTCATTGGGTTTCATAGAAA	5879
Dd	47762	TATGATGCCAAATGCTGCAAAAGATAAGAAATAAGATTCATTGGGTTTCATAGAAA	47703
OY	5880	AGTATATGGAAAAACATGTGTAAAAAAACATTGTGTAAATGACACATCTGTCAAAAGATT	5939
Dd	47702	AGTATATGGAAAAACATGTGTAAAAAAACATTGTGTAAATGACACATCTGTCAAAAGATT	47643
OY	5940	TTATATGGGGGATGAAATTTTGTATTTTCAAGAGACAAACAGTCCATACATAGCCAATCT	5999
Dd	47642	TTATATGGGGGATGAAATTTTGTATTTTCAAGAGACAAACAGTCCATACATAGCCAATCT	47585
OY	6000	AGTGTGTGAACCAAGGAGTTAGTGTCTGAAGTGGATTGGAGAAAGCATCATTTGAGCTG	6059
Dd	47582	AGTGTGTGAACCAAGGAGTTAGTGTCTGAAGTGGATTGGAGAAAGCATCATTTGAGCTG	47522
OY	6060	AGTTGGCTAGAGCTGTCTTCATAGACACTAATGTATATGATCAACACTGTGATCCAA	6119
Dd	47522	AGTTGGCTAGAGCTGTCTTCATAGAGCACTAATGTATGATGATCAACACTGTGATCCAA	47463

QY 6120 GTGCCCATCTCTCATGTAATGACAGAGGCGATTGAGATTCAGTGAATGACCGCTAA 6179
|||||
Db 47462 GTGCCCATCTCTCATGTAATGACAGAGGCGATTGAGATTCAGTGAATGACCGCTAA 47403
|||||
QY 6180 AGAAGATGATGAGATGTGG 6201
|||||
Db 47402 AGAAGATGATGAGATGTGG 47381
|||||
RESULT 4
AC026170 145212 bp DNA linear PRI 08-NOV-2002
LOCUS Homo sapiens chromosome 3 clone RP11-163D23 map 3p, complete
DEFINITION
sequence.
AC026170 145212 bp DNA linear PRI 08-NOV-2002
VERSION AC026170.5 GI:24796716
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Mu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,
Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,
Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C.,
Li, F., Li, G., Li, J., Li, S., Li, T., Liu, Y., Liu, N., Liu, B.,
Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, L., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X.,
Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B.,
Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.
TITLE
JOURNAL
REFERENCE
AUTHORS
Unpublished
Chromosome 3p genomic sequence
2 (bases 1 to 145212)
He, L., Hu, S., Dong, W., Wang, J., Zhang, Y., Zhang, H., Liu, B., Bao, W.,
Sun, Y., Wu, Q., Wang, H., Yang, X., Cheng, C., Wang, Y., Niu, Y., Qi, X.,
Li, T., Zhang, H., Liu, N., Wu, D., Yu, B., Fan, H., Liu, Y., Li, G.,
Li, C., Bao, Q., Bao, J., Wang, X., Song, L., Zhang, L., Guo, D.,
Huang, F., Zhang, G., Li, J., Bian, X., Zhang, M., Li, L., Feng, X., Yu, J.,
and Yang, H.
TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (21-MAR-2000) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
3 (bases 1 to 145212)
Mu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,
Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,
Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C.,
Li, F., Li, G., Li, J., Li, S., Li, T., Liu, Y., Liu, N., Liu, B.,
Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X.,
Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B.,
Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.
TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (19-JUL-2001) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing
100101, P.R.China
4 (bases 1 to 145212)
Mu, Q., Bao, J., Bao, Q., Bao, W., Bian, X., Cao, T., Chen, C., Chen, J.,
Ding, H., Dong, W., Fan, H., Feng, X., Gong, J., Guan, Q., Gu, X., Guo, D.,
Guo, Z., He, L., Hu, S., Huang, F., Jin, Y., Kang, N., Li, C., Li, C.,
Li, F., Li, G., Li, J., Li, S., Li, T., Liu, Y., Liu, N., Liu, B.,
Liu, Y., Li, W., Li, W., Li, Y., Luo, J., Niu, Y., Qi, Q., Qi, X., Song, L.,
Song, S., Sun, M., Sun, W., Sun, Y., Tan, X., Tao, R., Wang, H., Wang, J.,
Wang, J., Wang, L., Wang, L., Wang, R., Wang, X., Wang, X.,
Wang, Y., Wu, D., Xie, F., Xuan, Z., Xue, Y., Yan, C., Yang, X., Yu, B.,
Zeng, Y., Zhang, G., Zhang, H., Zhang, H., Zhang, L., Zhang, M.,
Zhang, X., Zhang, X., Zhang, Y., Zhang, Y., Zhang, Z., Zhu, B., Zhu, N.,
Yu, J. and Yang, H.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (08-NOV-2002) Human Genomic Center, Institute of
Genetics, Chinese Academy of Sciences, Datun Road, Beijing
100101, P.R.China
On Nov 8, 2002 this sequence version replaced gi:14916135.
-----Genome Center
Center: Beijing Center
Website: http://hg19p.ac.cn
http://www.genomics.org.cn
Contact: hgc@19p.ac.cn
-----Project Information
Center Project name: 1% project
Center Clone name: RP11-163D23
-----Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator; ET 5% of reads
Chemistry: Dye-terminator Big Dye; 45% of reads
Assembly program: PHRAP; version 0.990329
Consensus quality: 14584 bases at least Q40
Consensus quality: 147954 bases at least Q30
Consensus quality: 148116 bases at least Q20
Insert size: 145212; sum-of-ctrls
Quality coverage: 5.82x in Q20 bases; sum-of-ctrls

FEATURES
source
1.145212
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/map="3p"
/clone="RP11-163D23"
ORIGIN
Query Match 6.8%; Score 419.8; DB 9; Length 145212;
Best Local Similarity 52.3%; Pred. No. 2.1e-84;
Matches 1329; Conservative 0; Mismatches 1152; Indels 58; Gaps 16;
QY 449 TACTAATTTTGTACATTTGGAATATTAATGTTGTTAGCGCTGCGTCAAGCCT 508
|||||
Db 92524 TACTAATAATTAATAATAATAATAATAATTTTAAAAAGCGCGCGTGTCAAGCCT 92583
|||||
QY 509 GTAATCCCAAGCCCTTTGGAGGCGCAGAGGCGAGATCATAGTCAAGGATTTGAGACC 568
|||||
Db 92584 GTAATCCCAAGCCTTTGGAGGCGCAGAGGCGAGATCATAGTCAAGGATTTGAGACC 92643
|||||
QY 569 AGCTAGCCCAAGATGATGAAACCCGCTCTCTAATAAGATACAAAATTTAGCCAGGTGT 628
|||||
Db 92644 ATCTGCGCTAACAAGGTGAATCCATCTCTAATAATAATAATAATTTAGCCAGGTGT 92703
|||||
QY 629 GGTGGCGCACACCTGTATTAACCCAGCTGCTCGAGAGTGAAGCAGAGAAATTCCTTGAAC 688
|||||
Db 92704 GGTGGCGCGTGTGTATGATCCGCGCTACTCGAAGGCTGAGCAGAGAAATGCGTGAAC 92763
|||||
QY 689 CCAGAGGCGAGG-GTGCGAGCGAGCAAGATCAACCAATGACCT---GTAGCCTGATG 744
|||||
Db 92764 CCAGAGGCGAGGCTTACAGTGAAGCAAGATGCGCAGCTGACCTCGCTGCTGCGGCG 92823
|||||
QY 745 ACAGGCGCAAGCTCCAACTCAAAAAAATAAAAAAATAATATGTAATGTTGCTATTA 804
|||||
Db 92824 ACAGAGCGAAATCTCGCTCAAAAAAATAAAAAAATAATATGTAATGTTGCTATTA 92883
|||||
QY 805 ACAATAATTAATGAGCAGTGAAGCAAGAGTGTCTAAGACCTATATCAATGTAATCTAGAT 864
|||||
Db 92884 ATGTGATGATGATGAGGCGGATTTTAAATGGAATATATAGCAGGACCGGTGCTCA 92943
|||||
QY 865 ACTGTGGGAACA-TAATTAAGGAACCTTAACCAAGTCTGGAAGTAAGTTTGGTGA-- 920
|||||
Db 92944 TGCCTGTATTCACAGACCTTTGGAGAGCTGAGGCGGAGATCACTGAGTTGGAGTT 93003
|||||
QY 921 AAGGAATGTTGAGAGCAAGGTTAAAGAGAGTGAATAAAAAAATAATAATCAAGTTT 980
|||||

93004 CGAGACACGCTGACCAACATGGAAGAAACCCCATGCTACTAAATAATACAAATTTAGCCG 93063
981 AGCTGTGGAGATGGATAGGGAGCTTAAGAGAAATCAAAATGGAATGTTTCATG 1040
93064 GGCATGTGGCTGTATCTCAGCTACTGCGGAGGCTGAGGACGAGAACTGCTGAAC 93123
1041 TATGTTAAGACCCCTGTAGGTGAAGCACTATCATTTAT----- 1081
93124 TGGGAGGCGAAGTTGACAGCGGCGGAGATGTGCTGACCTTGGCTTGGCAACA 93183
1082 --CTGCACCATCGCGGAGCTTTTTTTTATGTGATGCTGGCAATTTAATAGAGAG 1139
93184 GGGTGAATCTCCGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 93243
1140 AGAGAAATGACAGTTGATGAGTCAAGATTGAAGTTCTGCGACAGATGTGAAGAA 1199
93244 ATTAACATTTCTTGGAGATTAAATGAGTAAACCTGGAAGACATTAATGACAGCC 93303
1200 GAGACAGTGAAGCAGAGAGTGAAGATTAATCAAGACAGAAATTAATGTGCGCAG 1259
93304 TGGTGTGAGACTTCAGTGCCTGAGTGAAGTCTGCTGCGGTTCAAAATTTGACCTGC 93363
1260 TGGCATCTGATCTGAGTCTAATCTGAGGAGAGATGAAGATTAACAGCTGCTGATAG 1319
93364 CATTTCTTATCTGTGACTTGAAGACTCAATTTAACTTCAATTTCTTCAATCCCTCA 93423
1320 TTATGAAGAGTGAAGGCTTCAAGAGCTTCAAGGTGTGAATTAATGAAGAAATGAT 1379
93424 TCTATATAGGAATG--TGATGATACTACTCTTGAAGTTGTAAGATTTAGTAAGTTT 93481
1380 GGAGAAAGATTAATCTGTGAAG--TGAGATTTTCAGGCTGTGATCTCAATTA 1435
93482 AGACGTAAAGTCTTGGAACAGTCCCTTCTCATATTAAGCAATGTAAGAGTTAGCTG 93541
1436 CCAGACACTGTCTAAATCTTCAAGAC-ATGATCCCTGCGCTCAAGGAGCTTACAGCC 1494
93542 TAAATTAATAGATATCTCCCTCTTTTCTAAATTAACACTGTGTTTCAAGGAAACCAA 93601
1495 AAAAACAAGATTAAGAAATACACACCAATTAATTAAGACACTTGTGTGAATATCA 1554
93602 GAAGCAATATTAATCTATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 93661
1555 AAAAAAATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1613
93662 GAAGCCACACTTGTGACCTTCTGATTAATCTTCTTCAAGCAATCCCGACTGATTTTG 93721
1614 -TCAAGAAATTTCAATGAGCGCGGCGGCTGAGCGCTGTAATCCAGACATTTG 1672
93722 CCCAGCATTTTACATTTTGAATTTAAGGTGCTCAAGCTGTATCCAGACATTTG 93781
1673 GAGGCGAGGCGGTGATCAAGAGTCAAGATCAAGACCAATCCCTGTGTTAACAAGTG 1732
93782 GAGGCGAGGCGGCGGATCAAGAGTCAAGAGTCAAGACCAATCCCGCTAACAAGTG 93841
1733 AAAACCCCTCTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1792
93842 AAAACCCCTCTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 93901
1793 TCCAGCTACTCAGAGGCTGAGGAGAGAAATGAGGTGAACCCCGGAGGTGAGATTG 1852
93902 TCCAGCTACTCAGAGGCTGAGGAGAGAAATGAGGTGAACCCCGGAGGCGGACTTGC 93961
1853 -GTGAGCCAGATTCGCGCACTGCGCTCCAGCTGCGGCAAGAGTGAAGTGCCTGCTCA 1911
93962 AGTGAAGCGAGATTCGCGCACTGCACTCCAGCTGCGGCAAGAGTGAAGTGCCTGCTCA 94021
1912 AAAAAAAAAAAAAAAAAAAG---TTTCAATCTTAACACACTGCTATTCAACAG 1967
94022 AAAAAAAAAAAAAAAAAAAGAAATTAATTAAGAACCTGTGCTCAATTTCAACTTTTCA 94081
1968 TCTAGAAATGAGAGATGTTTACAGGAGAGAAATGTTTTCAGAAAGTACAGAGTAG 2027
94082 GCTCACCTGGAAGATGGCTACTTCAATTTTAAAAAATTTATTTGAGGTGAATGACCA 94141

2028 GGAATTAAGCATATGTTCAAGAAAGAGACCCAGAGTCAATGCTTTGTTAGGTTAG 2087
94142 TTTTAACTGAACATTTAGTGGCATTTAGTACATTTTATTTTACAGCACACCCCTTA 94201
2088 GAAACACAGTGTGTTGCAATCTCCAGGTTCCATTAATGAGGTTATGAATCAATATG 2147
94202 GTTCAAAACATTTCCATTAATCTCAAAAGTAAACCCCTTACCATCAACAGGTTCTTC 94261
2148 TTAGCACTGCTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2207
94262 CCATCC-CTGAAGCTGCGCAATCACCAATATGATCTGTCACTATGATTTTCTAT 94320
2208 CATGCAATTAATTTGAAAGCAAGATTAATTTTCTGCACTTTGCTCCAA---TTG 2263
94321 TCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 94380
2264 TAAGTATCTTATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2323
94381 CACTGGCATTAATGTTTTCAGAGTTCATCCATGATGATGATGATGATGATGATGATGAT 94440
2324 GTAGACTGCAATCAAAAATC---TTTAAAGCACTGCGCTGCTTCACTTCTTATTTTG 2380
94441 TTTTCATGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 94500
2381 CAGAGAGAAATCCAGATCTGAGAGACAAACATTTTCTGAGGTTATGAACAGCT 2440
94501 CTGTTGAGGACATTTGGGCTGTTTCCATTTTGGGACCTGTGTGCTGAATTAATG 94560
2441 T-----ATGCCATTTGCTAAATGATTTCTAGTTAAATTTCTTCCACTAGTGCAT 2495
94561 TTTTACATGCTGGGTTGCTATCCCTTCTTATGATCTGCTGTTGGTGTGAGAG 94620
2496 TGCACCTTATGTTCTGTTGCTGCTGAATTAAGAAATTAATTAAGAAACAGCATACA 2555
94621 GTCAAGTGTCAAAAGTTATTAAGAAAGTGTCAAGAGTAAATGAAGAAAGATAG 94680
2556 TCTGGGAAATTAATTTGGTGAAGTGTGAGAGCTCATTTTCTAAGAAATGACCTT 2615
94681 TCTGAATGTGTTGAAGAAAGATGAAGATGATGATGATGATGATGATGATGATGAT 94740
2616 AGGCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2675
94741 AGGCGGCGAGTGTGCTCAAGCTGTTATTCACCACTTTGGAGGCTGAGGCGAG 94800
2676 ATGCTTGAACCCCAAGATTTCAAGACTAGCATGAGCAATTAAGCAAGACTGATCTTAC 2735
94801 ATCATCTGAGTCAAGAGTTTAAGACAGCTGCGCAACATGAACAAATCCATCTTAC 94860
2736 AAAAAATTTAAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2795
94861 TAAAAATCAAAAATTAAGCTGAGGATGATGATGATGATGATGATGATGATGATGAT 94920
2796 AGCTAGTGGGTGA--TCGCTGACAGAGATTTGAGGCTTAAGTGAAGCATATCAC 2853
94921 GGTGAAGCAAGATTTCTTGAACCGGAGGTTGAGAGTTGCAAGTGAAGCAAGTGT 94980
2854 ACAATGCACTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2913
94981 GGCATGCACTCAGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 95040
2914 TGGATTTTATCTTGAAC 2932
95041 AAAAAAGATGATCGTTGGC 95059

RESULT 5
AC026153 185321 bp DNA linear HTG 07-JUL-2000
LOCUS
DEFINITION Homo sapiens chromosome N/A clone Rpl1-269P18, WORKING DRAFT
SEQUENCE, 17 unordered pieces.
AC026153
AC026153.10 GI:8954068

KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 185321)
 Abola, A.P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Federspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R., Mao, J., Marathe, R., Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelm, J., Yu, S. and Davis, R.W.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 185321)
 Abola, A.P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Federspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R., Mao, J., Marathe, R., Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelm, J., Yu, S. and Davis, R.W.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2000) DNA Sequencing and Technology Center
 Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
 COMMENT On Jul 7, 2000 this sequence version replaced gi:8927681.
 ----- Genome Center
 Center: Stanford DNA Sequencing and Technology Development
 Center
 Center code: SDSTDC
 Web site: http://sequence-www.stanford.edu/group/human/
 Contact: hum-info@sequence.stanford.edu
 ----- Project Information
 Center project name: RPI1-269P18
 Center clone name: RPI1-269P18
 ----- Summary Statistics
 Sequencing Vector: M13mp18; X02513
 Chemistry: Dye-primer; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 175261 bases at least Q40
 Consensus quality: 179538 bases at least Q30
 Consensus quality: 180820 bases at least Q20
 Insert size: 178094; agarose-fp
 Insert size: 183721; sum-of-coverage
 Quality coverage: 8.2x in Q20 bases; sum-of-coverage
 Quality coverage: 8.0x in Q20 bases; sum-of-coverage
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 17 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1
 * 1657: contig of 1657 bp in length
 * 1658 1757: gap of unknown length
 * 1758 3546: contig of 1789 bp in length
 * 3547 3646: gap of unknown length
 * 3647 5277: contig of 1631 bp in length
 * 5278 5377: gap of unknown length
 * 5378 8396: contig of 3019 bp in length
 * 8397 8496: gap of unknown length
 * 8497 11351: contig of 2855 bp in length
 * 11352 11451: gap of unknown length
 * 11452 15441: contig of 3990 bp in length
 * 15442 15541: gap of unknown length
 * 15542 21194: contig of 5653 bp in length
 * 21195 21294: gap of unknown length
 * 21295 26148: contig of 4854 bp in length
 * 26149 26248: gap of unknown length
 * 26249 33667: contig of 7318 bp in length
 * 33668 33667: gap of unknown length
 * 33667 41541: contig of 7875 bp in length
 * 41542 41641: gap of unknown length
 * 41642 49652: contig of 8011 bp in length
 * 49653 49752: gap of unknown length
 * 49753 65062: contig of 15310 bp in length

FEATURES
 source
 1..185321
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="N/A"
 /clone="RPI1-269P18"
 /clone_lib="RPI1-269P18"
 /clone_id="RPI1-269P18"
 1..1657
 /note="assembly_name:Contig31"
 1758..3546
 /note="assembly_name:Contig35"
 3647..5277
 /note="assembly_name:Contig36"
 5378..8396
 /note="assembly_name:Contig37"
 8497..11351
 /note="assembly_name:Contig38"
 11452..15441
 /note="assembly_name:Contig39"
 15442..21194
 /note="assembly_name:Contig40"
 21295..26148
 /note="assembly_name:Contig41"
 clone_end:SP6
 vector_side:right
 26249..33667
 /note="assembly_name:Contig42"
 33667..41541
 /note="assembly_name:Contig43"
 41642..49652
 /note="assembly_name:Contig44"
 49753..65062
 /note="assembly_name:Contig45"
 clone_end:R7"
 65163..80491
 /note="assembly_name:Contig46"
 80592..101499
 /note="assembly_name:Contig47"
 101600..129616
 /note="assembly_name:Contig48"
 129717..152331
 /note="assembly_name:Contig49"
 152432..185321
 /note="assembly_name:Contig50"
 ORIGIN
 Query Match 6.8%; Score 419.8; DB 2; Length 185321;
 Best Local Similarity 52.3%; Pred. No. 2.1e-84;
 Matches 1329; Conservative 0; Mismatches 1152; Indels 58; Gaps 16;
 * 65063 65162: gap of unknown length
 * 65163 80491: contig of 15329 bp in length
 * 80492 80591: gap of unknown length
 * 80592 101499: contig of 20908 bp in length
 * 101500 101599: gap of unknown length
 * 101600 129616: contig of 28017 bp in length
 * 129617 129716: gap of unknown length
 * 129717 152331: contig of 22615 bp in length
 * 152332 152431: gap of unknown length
 * 152432 185321: contig of 32890 bp in length.
 Location/Qualifiers
 1..185321
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="N/A"
 /clone="RPI1-269P18"
 /clone_lib="RPI1-269P18"
 /clone_id="RPI1-269P18"
 1..1657
 /note="assembly_name:Contig31"
 1758..3546
 /note="assembly_name:Contig35"
 3647..5277
 /note="assembly_name:Contig36"
 5378..8396
 /note="assembly_name:Contig37"
 8497..11351
 /note="assembly_name:Contig38"
 11452..15441
 /note="assembly_name:Contig39"
 15442..21194
 /note="assembly_name:Contig40"
 21295..26148
 /note="assembly_name:Contig41"
 clone_end:SP6
 vector_side:right
 26249..33667
 /note="assembly_name:Contig42"
 33667..41541
 /note="assembly_name:Contig43"
 41642..49652
 /note="assembly_name:Contig44"
 49753..65062
 /note="assembly_name:Contig45"
 clone_end:R7"
 65163..80491
 /note="assembly_name:Contig46"
 80592..101499
 /note="assembly_name:Contig47"
 101600..129616
 /note="assembly_name:Contig48"
 129717..152331
 /note="assembly_name:Contig49"
 152432..185321
 /note="assembly_name:Contig50"
 Db 449 TACTAATTTTGGACCTTGAATATATAGTTGTTAGCGCTGCGGCTGCAAGCCT 508
 |||||
 53345 TACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 59404
 |||||
 Db 509 GTAATCCAGCCCTTTGGAGGCGCAGAGTGGCAATCATGAGTCAAGAGTTTGAAGACC 568
 |||||
 59405 GTATCCGACACTTTGGAGGCGCAGAGTGGCAATCATGAGTCAAGAGTTTGAAGACC 59464
 |||||
 Db 569 AGCTTAGCCAACTGTGAAACCCGCTCTCTAAGTACAAAAAATTAGCAGAGTGT 628
 |||||
 59465 ATCTGTGCTAACACGGGTGAATCCCTCTCTAATAAATAAATAAATAAATAAATAAATAA 59524
 |||||

Db 61682 TAAATAATACAAAAATTAGCTGGCATGTGCGGCATGCTTGTATGCCAGCTACTCTGTA 61741
QY 2796 AGCTAGTGGGTGGA--TCGCTTGACACAGAGTTTGAAGCTTAAGGTGAGCCATGATCAC 2853
Db 61742 GGCTGAGGACAAAGATTGCTTGAACCGGGAGATTGAGAGTTGCACTGAGCCAAATTTGT 61801
QY 2854 ACAATCTGACTCCAGCTTGAAGTACAGAGAGAACCTCTTCCCTTAAGAAAGAAAG 2913
Db 61802 GGCACAGCACTCCAGCTGGCGACAGAGAGAACCTCTGTCTCAAAAAAAGAAAAA 61861
QY 2914 TGGATTTATTCCTTAGAC 2932
Db 61862 AAAAAGATGATCCCTTGGC 61880

RESULT 6
HSM808294 2001 bp mRNA linear PRI 30-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp686G1080 (from clone DKFZp686G1080).
DEFINITION BX648147 GI:34367306
VERSION
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2001)
AUTHORS Ansoyge,W., Krieger,S., Regier,T., Rittmuller,C., Schwager,B.,
Mewes,H.W., Weill,B., Amid,C., Oeanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSTRM The German Human CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuerberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the CDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686G1080) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cdna/.
Location/Qualifiers
1..2001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686G1080"
/tissue_type="human endometrium carcinoma cell line"
/clone_id="686 (synonym: hlc03). Vector pSPORT1_Sfi; host
DH10B; sites SfiI + SfiIb"
/dev_stage="adult"
1956..1961
polyA_signal
1982
polyA_site
1982

Query Match 6.7%; Score 412.8; DB 9; Length 2001;
Best Local Similarity 92.1%; Pred. No. 7.7e-83;
Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 1 CACCATTGGGCTCAGGCAATTTATGAAAGCCAAATATACAGCTTAAATGAAATGTGAC 60
Db 53 CACCACGGGCTCGACAAATTTATGAAAGCCAAATATACAGCTTAAATGAAATGTGAC 112
QY 61 CTTAAATCCGAGAGAACTCCCTTTGTAGATTTGTAACAAATTAATTAATGAGTAG 120
Db 113 CTTAAATCCGAGAGAACTCCCTTTGTAGATTTGTAACAAATTAATTAATGAGTAG 172
QY 121 TTAATAGTTTAAATGAGATGTAACCCAGAGCCATATCAGCGCTAGCAAAATGGCAGA 180
Db 173 TAAATAGTTTAAATGAGATGTAACCCAGAGCCATATCAGCGCTAGCAAAATGGCAGA 232

QY 181 ATTCAATATCATCAAAAGTTATCCTTCAAGAGCTTACGCGCTTAATGATGCTTAAGAA 240
Db 223 ATTCAATATGACATCAAAAGTTGCTTCAAGAGCTTACGCGCTTAATGATGCTTAAGAA 292
QY 241 ATGTAAACGGCCCTCAGGCATCTGAAGACAGTGTATCAGCAATGATCAAAAAAGAAAA 300
Db 229 ATATAAACTCCCTTACGCGCCGCGAGACAGTGTATCAGCAATGATCAAAAAAGAAAA 352
QY 301 CCACAGGCCCTTCCCTTCCCATCTGATGTAAGACATCTTCAATTTCCATGTAGT 360
Db 353 CCACAGGCCCTTCCCTTCCCATCTGATGTAAGACATCTTCAATTTCCATGTAGT 412
QY 361 AAATTTTCTAGATACAGCTTGAAGCTCAAAAGTACGAAAGAGCTCCCATCAAG 420
Db 413 AAATTTTCTAGATACAGCTTGAAGCTCAAAAGTACGAAAGAGCTCCCATCAAG 472
QY 421 GAAATTTATCTTAAGATACGTAATGATTAATA-TTTTGTACATTTGGAATATATAAG 479
Db 473 GAAATTTATCTTAAGATACGTAATGATTAATA-TTTTGTCCATTTGGAATATATAAG 532
QY 480 TTGT 483
Db 533 TTGT 536

RESULT 7
HSM807657 2002 bp mRNA linear PRI 30-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp686I21199 (from clone DKFZp686I21199).
DEFINITION BX647511 GI:34366668
VERSION
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2002)
AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Weill,B., Amid,C., Oeanger,A.,
Fobo,G., Han,M. and Wiemann,S.
CONSTRM The German Human CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuerberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Duesseeldorf/Germany) within the CDNA
sequencing consortium of the German Genome Project. This clone
(DKFZp686I21199) is available at the RZPD in Berlin. Please contact
the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cdna/.
Location/Qualifiers
1..2002
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686I21199"
/tissue_type="human fetal brain"
/clone_id="686 (synonym: hlc03). Vector pSPORT1_Sfi; host
DH10B; sites SfiI + SfiIb"
/dev_stage="fetal"
1948..1953
polyA_signal
1975
polyA_site
1975

Query Match 6.7%; Score 412.8; DB 9; Length 2002;
Best Local Similarity 92.1%; Pred. No. 7.7e-83;
Matches 446; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 1 CACCATTGGGCTCAGGCAATTTATGAAAGCCAAATATACAGCTTAAATGAAATGTGAC 60

```

Db      47  CACCCAGGCGCTGGAGCAATTTATGAAACCGAATATACGCTTAAATATGAAATGGAGAC 106
Qy      61  CTAAATATCCAGAGAAAGTCCCTTGTGTAGATTTGTAAACAAAATTAATATGATGAGAG 120
Db      107  CTAAATATCCAGAGAAAGTCCCTTGTGTAGATTTGTAAACAAAATTAATATGATGAGAG 166
Qy      121  TTAATATGTTCTTAATGGAATGTTGAACCCAGAGCCATATCAGCGCTAGACAAAATGGCAGA 180
Db      167  TAAATAGTTCCTAATGAGTGTGTGAGCCAGAGCCATATCAGTGTAGGAAAATGGCAGA 226
Qy      181  ATTATATATATCATCAAGTATATCTTCACAGAGCTTGAGGCGCCAAAGATGCTAAAGAAA 240
Db      227  ATTCATATATGACATCAAGTGTGTCTTCAGAGAGCTTGCGCGCTTAAGATGTTCTAAAGAAA 286
Qy      241  ATGTGAAGAGCCCTCAGCATCTGAAGAGCAGTGTGTACAGCAATGTATGACAAAAGAAAA 300
Db      287  ATATGAAATCTCCCTCAGCGCGCCGAGAGACAGTGTATACAGCAATTAATCAAAAAGAAAA 346
Qy      301  CCACAGGCGCTTCCCTTCCCTCCCAATCTTGATGTAAAGCAGTCTTCAATTTTCCATAGT 360
Db      347  CCACAGGCGCTTCCCTTCCCTCCCAATCTTGATGTAAAGCAGTCTTCAATTTTCCACAGT 406
Qy      361  AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 420
Db      407  AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 466
Qy      421  GAAATTTATCTTAAGATCTGTAAATGATATCTAA-TTTTTGTACATTTGGAATATATTAAG 479
Db      467  GAAATTTATCTTAAGATCTGTAAATGATATCTAA-TTTTTGTCCATTTGGAATATATTAAG 526
Qy      480  TTGT 483
Db      527  TTGT 530

```

```

RESULT 8
BC008944      1846 bp      mRNA      linear      PRI 04-NOV-2003
LOCUS        Homo sapiens ubiquitin-conjugating enzyme E2 variant 1, mRNA (CDNA
DEFINITION   BC008944
ACCESSION   BC008944
VERSION      BC008944.2 GI:38014077
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1846)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Martusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Scapleton,M., Soares,M.B., Bonaldo,M.P., Caesavant,T.L.,
Schaeetz,T.E., Brownstein,M.J., Uedlin,T.B., Toshiyuki,S.,
Carinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,D.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalska,U., Smallue,D.E.,
Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
MEDLINE      22388257
PUBMED      12477932
REFERENCE     2 (bases 1 to 1846)

```

```

AUTHORS      Strausberg,R.
TITLE        Direct Submission
JOURNAL      Submitted (29-May-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA.
REMARK       NIH-MGC Project URL: http://mgc.nhl.nih.gov
COMMENT       On Oct 28, 2003 this sequence version replaced gi:14286289.
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting
FEATURES
source       1. 1846
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2960953"
/tissue_type="muscle, rhabdomyosarcoma"
/clone_id="NIH MGC_17"
/lab_host="DH10B-R"
/notes="vector: POTB7"
<1. 1846
/gene="UBE2V1"
/notes="synonyms: UBEV1, CIR1, CROC-1, UEV1A, CROC1, UEV-1"
/db_xref="locusid:7335"
/db_xref="MIM:602995"
<1. 431
/gene="UBE2V1"
/codon_start=3
/product="UBE2V1 protein"
/protein_id="AAH08944.2"
/db_xref="GI:38014078"
/db_xref="locusid:7335"
/db_xref="MIM:602995"
/translation="GSGVAVPRNPRLLBELBEGQKVGQGTGVSQGLDEEDMTLTRMT
GMIIQPPRIYENRISLKIICGPPYPEAPPEPVTKINNMGNVSNQGVDPRAISV
LAKMNSYSIKVYLOELRLMMSKMNMLPQPEGQCSN"
12. 428
/gene="UBE2V1"
/notes="OU con; Region: Ubiquitin-conjugating enzyme.
Protein destined for proteasome-mediated degradation may
be ubiquitinated. Ubiquitination follows conjugation of
ubiquitin to a conserved cysteine residue of UBC
homologues. TSG101 is one of several UBC homologues that
lacks this active site cysteine"
/db_xref="CD:pfam00179"
ORIGIN
Query Match 6 5%; Score 402.6; DB 9; Length 1846;
Best Local Similarity 92.5%; Pred. No. 16-80;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
16 CAATTATGAAGCAAAATATACACCTTAATATGATGTGACCTTAATATCCAGAG 75
160 CAATTATGAAGCAAAATATACACCTTAATATGATGTGACCTTAATATCCAGAG 219
Qy 76 AACTCCCTTTTGAAGATTGTGAACAAAATTAATATGATGTGATTAAGTTTATG 135
Db 220 CACCCCTTTTGAAGATTGTGAACAAAATTAATATGATGTGATTAAGTTTATG 279
Qy 136 GAATGTGAAGCAAGAGCATATCAGCGCTGAACAAAATGCGAATCATATATCATCA 195

```

Db 280 GAGTGTGAGACCCAGAGCCATATCACTGCTAGCAAAAATGCGAATTCATATAGCATCA 339

Qy 196 AAGTTATCCCTTCAAGAGCTTCAGCGCTATATGATGCTTAAAGAAATGTGAAACGCCCTC 255

Db 340 AAGTTATCCCTTCAAGAGCTTCAGCGCTATATGATGCTTAAAGAAATGTGAAACGCCCTC 399

Qy 256 AGCCATCTGAGAGACAGTGTATACAGCATTTGATCAAAAAAGAAACCAACGAGCCCTTCC 315

Db 400 AGCCGCGCCGAGAGACAGTGTATACAGCATTTAATCAAAAAAGAAACCAACGAGCCCTTCC 459

Qy 316 CTTCCGCCCACTGTTAGTATGAGAGCTTCATTTCCATTTAGTATGATTTTCTAGATAC 375

Db 460 CTTCCGCCCACTGTTAGTATGAGAGCTTCATTTCCATTTAGTATGATTTTCTAGATAC 519

Qy 376 AGCTTAGAGCTCAAAAGTACTGAAAAGAAAGCTCCCATTCAGAAATTTATCTTAAG 435

Db 520 GTCTTTAGAGCTCAAAAGTACTGAAAAGAAAGCTCCCATTCAGAAATTTATCTTAAG 579

Qy 436 ATACTGTAATGATCTAA-TTTTGTACTTTGGAATATATAGTTGT 483

Db 580 ATACTGTAATGATCTAA-TTTTGTACTTTGGAATATATAGTTGT 628

RESULT 9
BC000468 1870 bp mRNA linear PRI 12-NOV-2003
LOCUS Homo sapiens ubiquitin-conjugating enzyme E2 variant 1, mRNA (cDNA
DEFINITION clone MGC:8586 IMAGE:2961017), complete cds.
ACCESSION BC000468
VERSION BC000468.2 GI:38197161
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1870)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Frange, C., Kana, S.S., Loquellano, N.A., Peters, G.J.,
Abramsen, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultk, S.W.,
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boutard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Schnerk, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 1870)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2000) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT On Nov 6, 2003 this sequence version replaced gi:12653396.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgl.nih.gov
Akhter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooke, S.,
Dieckrich, N.L., Granite, S., Guan, X., Gupta, J., Hachighi, P.,
Hansen, N., Ho, S.-L., Karlins, B., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W.,
Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 1 Row: 1 Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10863894.

FEATURES
source
1..1870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:8586 IMAGE:2961017"
/tissue_type="Muscle, rhadomyosarcoma"
/clone_id="NTM_MGC_17"
/lab_host="DH10B-R"
/note="vector: pOTB7"
1..1870
/gene="UBE2V1"
/note="synonyms: UBEV1, CIR1, CROC-1, UBEVIA, CROC1, UEV-1"
/db_xref="LOCUSID:7335"
/db_xref="MIM:602995"
11..454
/gene="UBE2V1"
/codon_start=1
/product="UBE2V1 protein"
/protein_id="AAH00468.1"
/db_xref="GI:12653397"
/db_xref="LOCUSID:7335"
/db_xref="MIM:602995"
/translation="MAATGSGVKNRFRLLLEBQKGVGDGTVMGLDEDDMT
LRRWGMIGPRTYENRYSLEKPCPKPEAPPFRFYTKIMGNVNSNGVDP
RAISLAKQNSYSIKVYLQELRRLRMKMKKLPQPEGCYSN"
35..451
/gene="UBE2V1"
/note="UQ con; Region: Ubiquitin-conjugating enzyme.
Protein destined for proteasome-mediated degradation may
be ubiquitinated. Ubiquitination follows conjugation of
ubiquitin to a conserved cysteine residue of UBC
homologues. TSG101 is one of several UBC homologues that
lack this active site cysteine"
/db_xref="CDI:pfam00179"

ORIGIN
Query Match 5.5%; Score 402.6; DB 9; Length 1870;
Best Local Similarity 92.5%; Pred. No. 1.6e-80;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
Qy 16 CAATTATGAAAGCCAAATATACAGCTTAAATGATGAGACTTAATACCAAGAG 75
Db 183 CAATTATGAAAGCCAAATATACAGCTTAAATGATGAGACTTAATACCAAGAG 242
Qy 76 AACTCCCTTTGTAGATTGTGAACAAAATTAATGATGAGAGTAACTTAAAG 135
Db 243 CACCCCTTTGTAGATTGTGAACAAAATTAATGATGAGAGTAACTTAAAG 302
Qy 136 GAATGTGAAGCCAAAGCATATACAGCTTGAACAAAATGAGAGATTCATATCATCA 195
Db 303 GAGTGTGAGCCCAAGAGCATATACAGCTTGAACAAAATGAGAGATTCATATCATCA 362
Qy 196 AAGTTATCCCTTCAAGAGCTTCAGCGCTATATGATGCTTAAAGAAATGTGAAACGCCCTC 255

```

Db      363 AACGTTCTCGAAGAGCTTCGGCGCTTAATGATGCTTAAGAAATATATACTCCCTC 422
Qy      256 AGCCATCTGAAGAGAGCTTTTACAGCAATGATCAAAAAGAAAACACAGGCCCTTCCC 315
Db      423 AGCCGCCCGAAGAGAGCTTTTACAGCAATGATCAAAAAGAAAACACAGGCCCTTCCC 482
Qy      316 CTTCCCGCATCTGTTGATGAGAGCTTCTTATTTCCATGATGTAATTTTCTAGATAC 375
Db      483 CTTCCCGCATCTGTTGATGAGAGCTTCTTATTTCCATGATGTAATTTTCTAGATAC 542
Qy      376 AGCTTGAGAGCTCAAGTACGTAAGAGAGCTCCCATCAAGAGAAATTAATCTTAG 435
Db      543 GTCTTGAGAGCTCAAGTACGTAAGAGAGCTCCCATCAAGAGAAATTAATCTTAG 602
Qy      436 ATACTGTAATGATGACTAA-TTTTGTACATTTGGAATATATAGTTGT 483
Db      603 ATACTGTAATGATGACTAAATTTTGTGTCATTTGGAATATATAGTTGT 651

```

```

RESULT 10
LOCUS   AR000150      1930 bp      DNA      linear      PAT 04-DEC-1998
DEFINITION   Sequence 1 from patent US 5736331.
ACCESSION   AR000150
VERSION     AR000150.1   GI:3962681
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 1930)
AUTHORS     Lin,S.H. and Rothofsky,M.Lynn.
TITLE       Method for identifying nucleic acids encoding c-fos promoter
activating proteins
PATENT: US 5736331-A 1 07-APR-1998;
LOCATION/Qualifiers
FEATURES
source      1..1930
/mol_type="unknown"
/mol_type="unassigned DNA"

```

```

ORIGIN
Query Match      6.5%; Score 402.6; DB 6; Length 1930;
Best Local Similarity 92.5%; Pred. No. 1.6e-80;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Qy      16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGAGACTTAATACCGAGAAG 75
Db      245 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGAGACTTAATACCGAGAAG 304
Qy      76 AACTCCCTTTGTAAGATTGTGAACAAAATTAATAGATGAGATTATAGTTCTAATG 135
Db      305 CACCCCTTTGTAAGATTGTGAACAAAATTAATAGATGAGATTATAGTTCTAATG 364
Qy      136 GAATGTTGAACCCAAAGCATATCAGCGCTAGCAAAATGCGAATTCATATATCATCA 195
Db      365 GAGTGTGAGCCCAAGAGCATATCAGTGTAGCAAAATGCGAATTCATATATGATCA 424
Qy      196 AAGTTATCTTCAAGAGCTTCAGCGCTATGATGTCCTAAAGAAAATGTAAGAGCCCTC 255
Db      425 AAGTTATCTTCAAGAGCTTCAGCGCTATGATGTCCTAAAGAAAATGTAAGAGCCCTC 484
Qy      256 AGCCATCTGAAGAGAGCTTTACAGCAATGATCAAAAAGAAAACACAGGCCCTTCCC 315
Db      485 AGCCATCTGAAGAGAGCTTTACAGCAATGATCAAAAAGAAAACACAGGCCCTTCCC 544
Qy      316 CTTCCCGCATCTGTTGATGAGAGCTTCTTATTTCCATGATGTAATTTTCTAGATAC 375
Db      545 CTTCCCGCATCTGTTGATGAGAGCTTCTTATTTCCATGATGTAATTTTCTAGATAC 604
Qy      376 AGCTTGAGAGCTCAAGTACGTAAGAGAGCTCCCATCAAGAGAAATTAATCTTAG 435
Db      605 GTCTTGAGAGCTCAAGTACGTAAGAGAGCTCCCATCAAGAGAAATTAATCTTAG 664

```

```

Qy      436 ATACTGTAATGATGACTAA-TTTTGTACATTTGGAATATATAGTTGT 483
Db      665 ATACTGTAATGATGACTAAATTTTGTGTCATTTGGAATATATAGTTGT 713

RESULT 11
LOCUS   HSU939360      1982 bp      mRNA      linear      PRI 03-OCT-1997
DEFINITION   Homo sapiens DNA-binding protein (CROC-1A) mRNA, complete cds.
ACCESSION   U939360
VERSION     U939360.1   GI:1066079
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 1982)
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       CROC-1 encodes a protein which mediates transcriptional activation
of the human FOS promoter
Gene 195 (2), 141-149 (1997)
JOURNAL     PUBMED 9305758
MEDLINE     9305758
PUBMED      9305758
REFERENCE   2 (bases 1 to 1982)
AUTHORS     Lin,S.L.
TITLE       Direct Submission
SUBMITTED   (25-OCT-1995) Stanley L. Lin, Tumor Biology,
Schering-Plough Research Institute, 2015 Galloping Hill Road,
Kenilworth, NJ 07033, USA
LOCATION/Qualifiers
FEATURES
source      1..1982
/mol_type="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CROC-1A"
/tissue_type="Brain"
1..1982
/gene="CROC-1A"
70..582
/gene="CROC-1A"
/function="transcriptional activation of c-fos
proto-oncogene promoter"
/codon_start=1
/product="DNA-binding protein"
/protein_id="AAB72015.1"
/db_xref="GI:1066080"
/translation="MPGEVQASYLKSSQSLSDGRLPEPKFKGSKSPORLLEEL
ERGGQVGDGTVSMGLDEDMWTLFRMGMITGPRTYENRIVSLKIEGKVPAP
PFRFVYTKIMNGVNSNGVDPRAISVLAQNISIVLVQLRLMSEKMKLP
QPEEGQCYSN"
1959..1964
/gene="CROC-1A"

```

```

ORIGIN
Query Match      6.5%; Score 402.6; DB 9; Length 1982;
Best Local Similarity 92.5%; Pred. No. 1.6e-80;
Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;

Qy      16 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGAGACTTAATACCGAGAAG 75
Db      311 CAATTATGAAAGCCAAATATACAGCTTAAATAGATGAGACTTAATACCGAGAAG 370
Qy      76 AACTCCCTTTGTAAGATTGTGAACAAAATTAATAGATGAGATTATAGTTCTAATG 135
Db      371 CACCCCTTTGTAAGATTGTGAACAAAATTAATAGATGAGATTATAGTTCTAATG 430
Qy      136 GAATGTTGAACCCAAAGCATATCAGCGCTAGCAAAATGCGAATTCATATATCATCA 195
Db      431 GAGTGTGAGCCCAAGAGCATATCAGTGTAGCAAAATGCGAATTCATATATGATCA 490
Qy      196 AAGTTATCTTCAAGAGCTTCAGCGCTATGATGTCCTAAAGAAAATGTAAGAGCCCTC 255
Db      491 AAGTTATCTTCAAGAGCTTCAGCGCTATGATGTCCTAAAGAAAATGTAAGAGCCCTC 550

```


QY	256	AGCAGTCATGGAAGCACTGTTTACAGCAATTGATCAATAAAAGAAAAACACAGAGGCCCTTCCC	315
Db	551	AGCGCGCCGAAAGACATGTGTAACGCAATTAATCAAAAAGAAAAACACAGAGGCCCTTCCC	610
QY	316	CTTCCCGCCCAACTGATGTAGTAAGCAGCTCTTCATATTTTCCATGTAGTAATTTTCTAGATAC	375
Db	611	CTTCCCGCCCAATCGATTTAATGATGCTCTTCATATTTTCCAGAGTAGTAATTTTCTAGATAC	670
QY	376	AGCTTGTAGAGCTCAAAAGTACTGGAAAAAGAAAGCTCCCATTTCAAAAGAAATTTATCTTTAG	435
Db	671	GCTCTGTAGAGCTCAAAAGTACCGGAAAAAGAAAGCTCCCATTTCAAAAGAAATTTATCTTTAG	730
QY	436	ATACTGTAAATGATACCTAA--TTTTGTACATTTGGAATATATTAAGTGT	483
Db	731	ATACTGTAAATGATACCTAAATTTTGTTCATTTGAAATATATTAAGTGT	779
RESULT 12			
LOCUS	HSU39361	2135 bp	mRNA linear PRI 03-OCT-1997
DEFINITION	Homo sapiens DNA-binding protein (CROC-1B) mRNA, complete cds.		
ACCESSION	U39361		
VERSION	U39361.1	GI:1066081	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 2135)		
AUTHORS	Rochefsky,M.L. and Lin,S.L.		
TITLE	CROC-1 encodes a protein which mediates transcriptional activation of the human FOS promoter		
JOURNAL	Gene 195 (2), 141-149 (1997)		
MEDLINE	97449289		
PUBMED	9305758		
REFERENCE	2 (bases 1 to 2135)		
AUTHORS	Lin,S.L.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-OCT-1995) Stanley L Lin, Tumor Biology, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033, USA		
FEATURES	Location/Qualifiers		
source	1..2135		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CROC-1B"		
	/tissue_type="Brain"		
gene	1..2135		
	/gene="CROC-1B"		
CDS	70..735		
	/gene="CROC-1B"		
	/function="transcriptional activation of c-fos proto-oncogene promoter"		
	/codon_start=1		
	/product="DNA-binding protein"		
	/protein_id="AAB72016.1"		
	/db_xref="GI:1066082"		
	/translation="MAVYKFRTHSPALAEQLYPMECFVPCLLIRGTFPTNQHKSHTYFGLPRVVTLLQDMHVLIPRKHRIHVSPEHYFCITGVKVPNNRFLDEBESQKVGNGDTVMSGEDDEMTLITRWGKMIIGPRITVENRIYSLKIEGPKPEAPFVRFVATKINNGVSNNGVDPRAISVLAKWQNSYISIKVVLQELIRLIMSKENMKLFQPEEGQOQXSN"		
	polyA_signal		
	2112..2117		
	/gene="CROC-1B"		
ORIGIN			
	6.5%; Score 402.6; DB 9; Length 2135;		
	Best Local Similarity 92.5%; Pred. No. 1.6e-80;		
	Matches 434; Conservative 0; Mismatches 34; Indels 1; Gaps 1;		
QY	16	CAATTATGAAAGCCAAATATACAGCCTTAATATGAAATGTGAGCCTTAATATCCAGAG	75

[illegible]


```

Db      ||||| 618 AATTGTTCTCGACAGGCTTGGCGCTTAATGATGCTAAAGAAATATGAATCTCCCTC 677
Qy      ||||| 256 AGCCATCTGAGAGACAGTGTATACAGCAATTCATCAAAAAGAAAACACAGCCCTCC 315
Db      ||||| 678 AGCCGCCCGAGAGACAGTGTATACAGCAATTCATCAAAAAGAAAACACAGCCCTCC 737
Qy      ||||| 316 CTTCCCTCCATCTGTAATGTAAGACAGTCTTCTTCATTCATGTAATTTTCTAGATAC 375
Db      ||||| 738 CTTCCCTCCCGAGAGCTTAATGATGCTTCTTCATTCATGTAATTTTCTAGATAC 797
Qy      ||||| 376 AGCTTGAGAGCTCAAGTACTGAGAAAGAAAGCTCCCATTCAGAAAGAAATTTATCTTAAG 435
Db      ||||| 798 GTTGTGAGACCTCAAGTACCGAAAGAGAGCTCCCATTCAGAAAGAAATTTATCTTAAG 857
Qy      ||||| 436 ATACTGTAATGATGTAAT-TTTTGTACATTTGTAATATATAGTTGT 483
Db      ||||| 858 ATACTGTAATGATGTAATTTTGTCCATTTGAATATATAGTTGT 906

```

```

RESULT 14
AC148906 229392 bp DNA linear HTG 21-APR-2004
LOCUS Ocolemur garnettii clone CH256-538H20, WORKING DRAFT SEQUENCE.
DEFINITION AC148906
AC148906.1 GI:46430802
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS Ocolemur garnettii (small-eared galago)
SOURCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM Mammalia; Eutheria; Primates; Strepsirrhini; Galagonidae; Ocolemur.

```

```

REFERENCE
AUTHORS 1 (bases 1 to 229392)
Antoniellis, A., Ayele, K., Benjamin, B., Blakeley, R.W.,
Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
Coleman, H., Daki, N., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B.,
Idol, J.R., Jones, C., Karlins, E., Kim, H., Kwong, P., Latic, P.,
Larson, S., Lee-Lin, S.-Q., Legaapi, R., Maduro, Q.L., Maduro, V.B.,
Marilyn, E.H., Mastello, C., Maskeri, B., McDowell, J.,
Mullikin, J.C., Paguirigan, C., Portnoy, M.E., Prasad, A., Puri, O.,
Reddy-Bugne, N., Schandler, K., Schueler, M.G., Shah, K., Sison, C.,
Stardrup, S., Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L.,
Wetherby, K.D., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 229392)
Green, E.D.
Direct Submission
Submitted (21-APR-2004) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: SPY
Center clone name: 538H20
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 229306 bases at least Q40
Consensus quality: 229348 bases at least Q30
Consensus quality: 229382 bases at least Q20
Insert size: 247000; agarose-fp
Insert size: 229392; sum-of-contigs
Quality coverage: 10.87x in Q20 bases; sum-of-contigs
Quality coverage: 11.70x in Q20 bases; sum-of-contigs

```

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces

```

FEATURES
source
1..229392
/organism="Ocolemur garnettii"
/mol_type="genomic DNA"
/db_xref="taxon:10611"
/clone="CH256-538H20"
/clone_1lb="CH256"
/notes="BAC resource: http://bacpac.chori.org/"
1..229392
/notes="assembly_fragment"
clone_end:SP6
vector_side:right"

```

```

ORIGIN
Query Match 6.2%; Score 387.2; DB 2; Length 229392;
Best Local Similarity 91.5%; Pred. No. 5.4e-77;
Matches 443; Conservative 0; Mismatches 38; Indels 3; Gaps 3;

```

```

Qy      1 CACCATGGGCTCAGGCAATTTATGAAGCAAAATATACAGCTTAAATAGAAATGGAC 60
Db      226746 CACCATGGGCTCAGGCAATTTATGAAGCAAAATATACAGCTTAAATAGAAATGGAC 226687
Qy      61 CTAATATCCCAAGAAAGCTCCCTTTGTAAGTTTGTACAAATATATATAGAGAGAG 120
Db      226686 CTAATATCCCAAGAAAGCTCCCTTTGTAAGTTTGTACAAATATATATAGAGAGAG 226627
Qy      121 TTAATAGTTCTAATGATGTAAGTGAACCCAGAGCATATCAGCGCTAGCAAAATGGCAGA 180
Db      226626 TTAATAGTTCTAATGATGTAAGTGAACCCAGAGCATATCAGCGCTAGCAAAATGGCAGA 226567
Qy      181 ATTCAATATATCAATCAAGTTATCTTCAAGAGCTTACGCGCTTAATGATGTTAAGAAA 240
Db      226566 ATTCAATATATCAATCAAGTTATCTTCAAGAGCTTACGCGCTTAATGATGTTAAGAAA 226507
Qy      241 ATGTGAAGCGCTCAGCATCTGAAGGACAGTGTATACAGCAATGATCAAAAGAAA 300
Db      226506 ATGTGAAGCGCTCAGCATCTGAAGGACAGTGTATACAGCAATGATCAAAAGAAA 226448
Qy      301 CCACAGGCGCTTCCCTTCCCATATCTGATGAAGCAGTCTTCAATTTCCATAGTAGT 360
Db      226447 CCACAGGCGCTTCCCTTCCCATATCTGATGAAGCAGTCTTCAATTTCCATAGTAGT 226389
Qy      361 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTACGAAAGAAAGAGCTCCCATTCGAAG 420
Db      226388 AAATTTTCTAGATACAGCTTGTAGAGCTCAAGTACGAAAGAGAGCTCCCATTCGAAG 226329
Qy      421 GAAATTTATCTAAGATCTGTAATGATGATCA-ATTTTGTACATTTGGAATATATAG 479
Db      226328 GAAATTTATCTAAGATCTGTAATGATGATCAATTTTGTACATTTGGAATATATAG 226269
Qy      480 TTGT 483
Db      226268 TTGT 226265

```

```

RESULT 15
AC008736 191925 bp DNA linear PRI 27-SEP-2000
LOCUS Homo sapiens chromosome 19 clone CTD-2538C1, complete sequence.
DEFINITION AC008736
AC008736.6 GI:10312244
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
ORGANISM Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

REFERENCE 1 (bases 1 to 191925)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 191925)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 191925)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Sep 27, 2000 this sequence version replaced gi:8575905.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.1.
STS Content:
SHGC-57769 G37408.
FEATURES
source Location/Qualifiers
1..191925
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2538C1"
ORIGIN
Query Match 6.2%; Score 386.6; DB 9; Length 191925;
Best Local Similarity 51.4%; Pred. No. 7.4e-77;
Matches 1266; Conservative 0; Mismatches 1144; Indels 53; Gaps 14;
QY 471 ATATATTAAGTTGTTAGCTGCGCGGTGGCTCACGCTGTATCCAGCCCTTTGGGAGG 530
DB 74762 AATATACGAGTTTACGCGGCTGTGGCTCACACCTGTATCCAGCCTTTGGGAGG 74821
QY 531 CCAAGTGGGAGATCATGAGTGAAGGTTGAGCAGGCTGACCAACATGATGTAAC 590
DB 74822 CCGAGTGGGAGATCAGAGGTGAGAGATGAGATCCATCTTACATACAGATGAAC 74881
QY 591 CCCGCTCTACTAAAGATCAAAAATTAGCCAGGTGTGTGGCGCACCTGTAACCC 650
DB 74882 TCCGCTCTACTAAAGATCAAAAATTAGCTGGGCGGTGTGGCGACATCTGTAGTCCC 74941
QY 651 AGCTGCTGAGAGATGAGGAGAGATTTGCTTGAACCCAGAGGAGAG-GTGCAAGG 709
DB 74942 AGCTACTGGGAGGCTGAGGAGAGAGATGTGAAACCGGAGTGTGGAGCTTGAGTG 75001
QY 710 AGCAAGATCAACCAATGACTGTAGGCTGATGACAGGCAAGACTCCAATCA-AAA 768
DB 75002 AGCCAGATGCGGCACTGCACTCCAGCTGGGCGACAGACCAAGACTCCATCCACAAA 75061
QY 769 AAAAAAAAAAAAAAATATGTAAAGTGTGCTATTAACAATAATAGGCGATGAAGC 828
DB 75062 AAAAAAAAAAAAAAATATGTAAAGTGTGCTATTAACAATAATAGGCGATGAAGC 828
QY 829 AAGAGCTTAAGCCATATGACATGTGTAACCTAGGAATCTGTGGGAACATATATAGGGA 888
DB 75122 AATATACCCAGATATGGAATTAACAGCGAGGACTTAAAAACAGTATATTAATACGA 75181
QY 889 ACCTAACCCAGTCTGAGAGTAAAGGTTTTGAAAGAAATGTTGAGCAAAAGGTTAA 948
DB 75182 ACTTAAGATGAGATGTAGTGAAGTGAAGTGAAGAAACAAGAAATCTCACTAAAGATGA 75241
QY 949 GAGAGTGAAGAAAAATTAATAATACAGTTTAACTGTGTGAGAAATGGAGATAGGAGACT 1008
DB 75242 AACTCATATGTATATAGATATGTATCTATTTATACATGTATATATATAGATACACAC 75301
QY 1009 AACTGAGAAATCAATATAGAAATGTTTATCATGTATGTTAAGAACCTGTGTAAGGTGAAG 1068
75302 ACAAGCTGTGAAAAACCTAAATTTCAATATGAAAAATTTAGGCCAGG----- 75355
QY 1069 ACCATTAATATATGTGACCATGCGGAGCTTTTATTTATGTTAAATGCTTGGCAATTTA 1128
DB 75356 -----CACGGTGGCTCATGCTCTGTAAATCCAGACACTTTGGAGGCGGAGCGGATC 75410
QY 1129 AATAGAGAGAGAGAAATGTAGACAGTTGATGATGATGATGATGATGATGATGATGATG 1188
DB 75411 ACAGAGTACAGAGATTCAGACCAATCTTAAGTAAACAGATAAATCTCGGTCTTAATAA 75470
QY 1189 TGTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1248
DB 75471 ATACAAAAATTAATCTGGGCTGTGTGACAGATACCCGTATCCAGATCTCGGAGAGCT 75530
QY 1249 GTGCTGGCCAGTGGCATCTACTGTAGTCTAATCTGAGGAGAGAGAGAGAGAGAGAGAG 1308
DB 75531 GAGGAGAGAGAAATGTGTGAAACCGGGTGTGTGAGCTTGCAGTGAAGCCGAGATTGCCCA 75590
QY 1309 CTGCTGATATGTTATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1368
DB 75591 CTGCACTCCAGCTTGGGCGACAGAGCAAGACTCATCTCACAAAAAATTTTCAAAAAA 75650
QY 1369 GAAGATGATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1422
DB 75651 AAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 75710
QY 1423 TGACTCTCAATACAGACAGACTGTGTAAATGCTTCAAGACATGATCCCTGCCCTCAAG 1482
DB 75711 TTACAGCGAGAGACTTTAAAAACAGCTATTTAAATACAGACTTAAAGATGAGATGTT 75770
QY 1483 GGAATTAACAGCCAAAAACAGAGATTAAGAAATACACCAATATATATAGACACTTG 1542
DB 75771 AAGAGATGTAAGAAACAAAGATCTCACTTAAGAAATGAAATCTATATATATATGA 75830
QY 1543 TGTAGAAATCAAGAAAGAAATACAGTCTAGTACTGTATGATGTGCAACGCGATCAAGAT 1602
DB 75831 TATGATATCATTTATATACATGTATATATATATATATATATATATATATATATAT 75890
QY 1603 ATCTTCTAGTTTCAAGAAATTTCAATGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1662
DB 75891 AATTTCTATTAATGAAATTTAAATTTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 75950
QY 1663 AGACCTTGGAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1722
DB 75951 AGACCTTGGAGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 76010
QY 1723 TTAACAGGTGAATCCCGCTCTAC--AAAAATTTAAATTTAGCCAGGCGTGTGGGC 1780
DB 76011 TTAACATGTGAATCCCATCTGTACAAAAAATTTAGCCAGGCGTGTGGGC 76070
QY 1781 GGGCGGCTGTATGTCAGGACTCTCAAGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1840
DB 76071 GGGCACCTTATAGTCCAGGACTCTCAAGAGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 76130
QY 1841 GGTAGAGTTTGC-GTGAGCGAGATGCGCACTGCGCTCCAGCTGGGCGAC-AGAGTG 1898
DB 76131 GGGGAGGCTTGAAGTGAACCAAGATGCGCCACTGCACTCCAGCTTGGGCGCAAGAGTG 76190
QY 1899 AGACTGCGTCAAAAAAATTTTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1954
DB 76191 AGACTGCTCAAAAAAATTTTCAAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 76250
QY 1955 TGCAATTTCAACAGCTAGAGATGAGAGAGATGTTACAGGAGAGAGAGAGAGAGAGAGAG 2014
DB 76251 GCTTAACATTTGTTAAAGTCCAGTATGATGATGATGATGATGATGATGATGATGATGAT 76310
QY 2015 AGGTACAGAGTGAAGAAATAGAGATATGTTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2074
DB 76311 GATTTAGAAAGATTAATTAATTTGCTTTATTTATTTATTTATTTATTTATTTATTTAT 76370
QY 2075 GTTAGGTTAGAGAAACACAGTGTTTGCATTCCTCAGGTTCCATTAAGTGTGCTATGAA 2134

```

DB 76371 GAACCTTAGAGCTCTATTAATAAATTAGACCTACATGAAAATTATAAAAGCTGCTGA 76430
QY 2135 ATCAATATGGTGGTTAGCACTGCAATTTTAAATAAATGAATGAATGAGAGAA 2194
DB 76431 AAGAAATTAAGAGACATTAATAATGAGAAAGACATCCACATTCATTGACTGAAAGAT 76490
QY 2195 TAGAAAATTTAGCATGCACTTACATTTTGAAGAGCAAGATATATTTCTGCAACTTTG 2254
DB 76491 TTAATATCTTAAGATGTCATCTACTCAAGCAATCTACACATTCATTCATCCCTA 76550
QY 2255 CTCCAATGTAATCTGACTTAATTTTATGATGATGAAATACAGATACATATATA 2314
DB 76551 CCAAAATCCACATGATATTTTTCAGAAATAGAAAAGCAGATTTCTAAATTCACATGAA 76610
QY 2315 TTTCTTACTGT---AGACTGAGTCAGAAAATCTTTAAAGCCTGCGCTGTACTT 2370
DB 76611 ATCCCAAGGAGCCCAAAATAGCCAAAACATATCAAAAAGAACAAAGCTGGAGACTCA 76670
QY 2371 CCTTATTTGCAAGAGAAATCCAGATCTGAGAGACAAACATTTTGCCTGAGGTTAT 2430
DB 76671 CACTTGATCACAAACCTTTACTCAAAAGCTACAGTAATCAAAACAGTATGTACTGAAAT 76730
QY 2431 AGAACCAAGCTTATGCCATGTCTAAAGTATCTTAGTTAAATCTTCCCACTAGTGC 2490
DB 76731 AAAGACAGCATATAGACCAACAGAGTAGAATAGAGATCCAGATTAACCTTGTGCTC 76790
QY 2491 CATACTGCACTTCTAGTCTGTGGCTGAGAAATACAGAAATATATTAAGTGAACAGCATAC 2550
DB 76791 TA-----GTCAATGATTTTGTGACAAAGGGTACCAAAACATTT-----CAGGG 76832
QY 2551 ACAAGCTGGGGAATATATTTGGGTAGTGGCTGAGAGCTCTCATTTTCTAAGAAATGTGG 2610
DB 76832 AAAGGTTTGTAACTAATAGACATGGGAAAATCTGATATTCACATGCAAAAGAAATG-AA 76891
QY 2611 ACCTTAGCAGGGTATGGTGGCTCACACCTATATTCAGACCTTGGAGGCGCAAGTCA 2670
DB 76892 CTTGAGGCGCAACAAATGGCTCACACCTGTAATCCAGCACTTGGAAAGGCCAAAGCA 76951
QY 2671 AGAAGATCGCTTGAACCCAGAGTTCAAGACTAGCATGGGCAACATAGCAAGACCTCATC 2730
DB 76952 GGCAGATCACCTGAGGTCAGAGTTGACACAAAGCTGGCCAAATGTAATCCCGTC 77011
QY 2731 TCTACAAAATTTTAAAAATCAGCTGAGCATGCTGCTAGTCCCACTACC 2790
DB 77012 TCTACTTAAAAA-TACAAAATTTAGCCGGGTGTGTGTACATGCCCGTAATCCAGCTACT 77070
QY 2791 TGGGAAGCTAGGTGGGTGATGCTTGACACA--GGAGTTTGAGGCTAAGGTAGGCCATG 2848
DB 77071 CAGGAGGCTACGCCAAGAGATCACTTGAACCTGGAGGTGAGGTTGCAGTAGCCGAA 77130
QY 2849 ATCAGACAACTGCACTCCAGCTTGAAGTGAAGAGAAAGCCCTGTCCCTTAAAAAGAAAG 2908
DB 77131 ATTGACACCACTGCACTCCGCGCTGGGTAAAGAGTGAAGCTGTCTCAAAAAAGAAATA 77190
QY 2909 AAA 2911
DB 77191 AAA 77193

Search completed: December 27, 2004, 04:18:15
Job time : 16943.5 secs

This Page Blank (uspto)

|||||
Db 241 CCTGTGTATATAAATAATAGCTGTGTCTTGGACAAGGCTCGAGCTACT 300
QY 301 CGGAACACTGAGGTGGAGGATCACTGAGCCAGAGGCGTGAAGCTGACAGTGA 360
Db 301 CGGAACACTGAGGTGGAGGATCACTGAGCCAGAGGCGTGAAGCTGACAGTGA 360
QY 361 TCACCCAGCTGATTCAGCTCTGGAAGACAGAGGAGACCTGTCTTCAAAAAA 420
Db 361 TCACCCAGCTGATTCAGCTCTGGAAGACAGAGGAGACCTGTCTTCAAAAAA 420
QY 421 AAAAAAATGCAAGAAAGACATCAATACTGACCTGGACATTAATTTATGTA 480
Db 421 AAAAAAATGCAAGAAAGACATCAATACTGACCTGGACATTAATTTATGTA 480
QY 481 AAATTCACATCTTTTAGAGAAATTAAGATTCTGATTAATATATATATAT 540
Db 481 AAATTCACATCTTTTAGAGAAATTAAGATTCTGATTAATATATATATAT 540
QY 541 ATTATTAATTCAAATGGAATTAATAATTCAGAAACTAGCTTCTCACTCTCA 600
Db 541 ATTATTAATTCAAATGGAATTAATAATTCAGAAACTAGCTTCTCACTCTCA 600
QY 601 CAGTCAAAAATTTAATGTCTTTGGCCGGGTGGGTGGCTCAGGCTGTAAATCC 660
Db 601 CAGTCAAAAATTTAATGTCTTTGGCCGGGTGGGTGGCTCAGGCTGTAAATCC 660
QY 661 TTGGAGAGCGCAGCGGAGTGATCAAGAAGTTAGAGATCGAGACATCTGGCTA 720
Db 661 TTGGAGAGCGCAGCGGAGTGATCAAGAAGTTAGAGATCGAGACATCTGGCTA 720
QY 721 CGGTGAACCTCGTCTCTACTAATAAATACAAAAAATTAAGCCGGGTGGGTG 780
Db 721 CGGTGAACCTCGTCTCTACTAATAAATACAAAAAATTAAGCCGGGTGGGTG 780
QY 781 TGTAGTCCACGCTGCTAGAGGCTGAGCAGAGAGATGTGTGAACCCGGAGGCG 840
Db 781 TGTAGTCCACGCTGCTAGAGGCTGAGCAGAGAGATGTGTGAACCCGGAGGCG 840
QY 841 CTTCGAGTGAAGCGAGATTCGCGCACTGCACTCAGCTGGGCGACAGTGCAG 900
Db 841 CTTCGAGTGAAGCGAGATTCGCGCACTGCACTCAGCTGGGCGACAGTGCAG 900
QY 901 TCTCAAAAAAATTTAATGTCTTTAGAGCCAAAGTATCTTCTTTC 960
Db 901 TCTCAAAAAAATTTAATGTCTTTAGAGCCAAAGTATCTTCTTTC 960
QY 961 TTTCTCTCTTTTCTTTTCTTTTCAAAAAATATCTAGATTGAATCTTGAAT 1020
Db 961 TTTCTCTCTTTTCTTTTCTTTTCAAAAAATATCTAGATTGAATCTTGAAT 1020
QY 1021 TAACTCTCTTCTTGTGGCAATTTGAAATGAAAAATATCATGCTCATATTA 1080
Db 1021 TAACTCTCTTCTTGTGGCAATTTGAAATGAAAAATATCATGCTCATATTA 1080
QY 1081 CTGAACATTTTAAAAAACAATCATGAGGTTCAAAATCAAAATCTCAAAAT 1140
Db 1081 CTGAACATTTTAAAAAACAATCATGAGGTTCAAAATCAAAATCTCAAAAT 1140
QY 1141 GATATATGACATACTTATATTTTCCCTTAATATGATTTGTTATATATCCCA 1200
Db 1141 GATATATGACATACTTATATTTTCCCTTAATATGATTTGTTATATATCCCA 1200
QY 1201 TGTCTCACTTATATATATATATATATATATATATATATATATATATAT 1260
Db 1201 TGTCTCACTTATATATATATATATATATATATATATATATATATATAT 1260
QY 1261 AGTGTGATTTGAAGCTACAGCTGAAACCAATTTAGCCAGATTTCTTGACAT 1320
Db 1261 AGTGTGATTTGAAGCTACAGCTGAAACCAATTTAGCCAGATTTCTTGACAT 1320
QY 1321 GTGCTGTTAATATAGTACTTATACATATATATATATATATATATATATAT 1380
Db 1321 GTGCTGTTAATATAGTACTTATACATATATATATATATATATATATATAT 1380

Db 1381 GTGCTGTTAATATAGTACTTATACATATATATATATATATATATATATAT 1380
QY 1381 CAAGCAAGAAAGAAAGAAATTTGATTTCAAACTCTCTCTTCATCACCATTGG 1440
Db 1381 CAAGCAAGAAAGAAAGAAATTTGATTTCAAACTCTCTCTTCATCACCATTGG 1440
QY 1441 ATATCATATTTTGTACATTTAAGAACAAATAGGAGCTCACACATATTTTGA 1500
Db 1441 ATATCATATTTTGTACATTTAAGAACAAATAGGAGCTCACACATATTTTGA 1500
QY 1501 TGAATGAATGGCAACCTTCTAAGACTATGGAATCACTATTTGTTGAAGCA 1560
Db 1501 TGAATGAATGGCAACCTTCTAAGACTATGGAATCACTATTTGTTGAAGCA 1560
QY 1561 GCAATGATATTTTCACTTTTCTGTGTTTATATCTGTGTCTTTGACATCT 1620
Db 1561 GCAATGATATTTTCACTTTTCTGTGTTTATATCTGTGTCTTTGACATCT 1620
QY 1621 AAGTTAGCTAGGTACCAATTTGTTTAAAGCTGAAAAACCAATGCTTCA 1680
Db 1621 AAGTTAGCTAGGTACCAATTTGTTTAAAGCTGAAAAACCAATGCTTCA 1680
QY 1681 GGTAGGAGAAAAATATCTCTCAACATGTCCAATTAGACCAAGAAAACTA 1740
Db 1681 GGTAGGAGAAAAATATCTCTCAACATGTCCAATTAGACCAAGAAAACTA 1740
QY 1741 TATCAACATCAATGATATATATATATATATATATATATATATATATATAT 1800
Db 1741 TATCAACATCAATGATATATATATATATATATATATATATATATATATAT 1800
QY 1801 AAACATATATATCAATATGATATATATATATATATATATATATATATAT 1860
Db 1801 AAACATATATATCAATATGATATATATATATATATATATATATATATAT 1860
QY 1861 TTTTATTTTATTAATTAACAGCTGCTGTTTCAATGTTGCAATCTGAATGA 1920
Db 1861 TTTTATTTTATTAATTAACAGCTGCTGTTTCAATGTTGCAATCTGAATGA 1920
QY 1921 TCCCTGTCTGATGTGAGTTTCTTGCATGATGAAGAAAACTGTAAGTTG 1980
Db 1921 TCCCTGTCTGATGTGAGTTTCTTGCATGATGAAGAAAACTGTAAGTTG 1980
QY 1981 GCTCCAGCAGAGCAGATCATGTGATCATATGAAGCTCCAGCTGTGACCT 2040
Db 1981 GCTCCAGCAGAGCAGATCATGTGATCATATGAAGCTCCAGCTGTGACCT 2040
QY 2041 AAAGGAGAGAAACAAGATAGAGAGCGATGGGGAAAGTTCAAGTCCGGAT 2100
Db 2041 AAAGGAGAGAAACAAGATAGAGAGCGATGGGGAAAGTTCAAGTCCGGAT 2100
QY 2101 CTTGAACCTTAAAGATTAAGGTTCAAGAGCTGTGCAAAAGACTGACCTGA 2160
Db 2101 CTTGAACCTTAAAGATTAAGGTTCAAGAGCTGTGCAAAAGACTGACCTGA 2160
QY 2161 GCTGCTACATTTGTTCAAGAGGAGGAGGAGCAACAAGCTGTGAGAGATCA 2220
Db 2161 GCTGCTACATTTGTTCAAGAGGAGGAGGAGGAGCAACAAGCTGTGAGAG 2220
QY 2221 CCAAGAGGCTGCTTCAAGATTTGTATCCCAAGAGAGACATCTTCTAT 2280
Db 2221 CCAAGAGGCTGCTTCAAGATTTGTATCCCAAGAGAGACATCTTCTAT 2280
QY 2281 AACCAAGAGGTTCAAGCACTCCCTTTTGTAGCGGATCTGATCTTCTGCG 2340
Db 2281 AACCAAGAGGTTCAAGCACTCCCTTTTGTAGCGGATCTGATCTTCTGCG 2340
QY 2341 AAACCAATTAATTAAGAAATTTATTAAGTCAAGAAATTTATGCTGTAT 2400
Db 2341 AAACCAATTAATTAAGAAATTTATTAAGTCAAGAAATTTATGCTGTAT 2400
QY 2401 TTTGGGAAATTTCTTGTAAACCAAAAGGAGAAATTAATCTTGGCTTGG 2460
Db 2401 TTTGGGAAATTTCTTGTAAACCAAAAGGAGAAATTAATCTTGGCTTGG 2460

Qy	2463	ACTCACTGGCTTGAAGTCGAGAAAGTGTCTCTCAAAATCTCTAAGGCTCTAAATTAC	2550
Dp	2461	ACTCACTTGGCTTGAAGTCGAGAAAGTGTCTCTCAAAATCTCTAAGGCTCTAAATTAC	2520
Qy	2521	AGACCTGAAACTTAAAGGCAAGCTGCAATATGATGGTATGCTATGGATTTGAACTT	2580
Dp	2521	AGACCTGAAACTTAAAGGCAAGCTGCAATATGATGGTATGCTATGGATTTGAACTT	2580
Qy	2581	TAGTAATTAAGTTCATGATTTATTACCAATGCAATGATTTATCCCTACAGCAATAAATTA	2640
Dp	2581	TAGTAATTAAGTTCATGATTTATTACCAATGCAATGATTTATCCCTACAGCAATAAATTA	2640
Qy	2641	AGTGAACATGAAAAAAAAGCCAGACTTAAACAGAAAAAAGTTGCAAAAATCATCAACA	2700
Dp	2641	AGTGAACATGAAAAAAAAGCCAGACTTAAACAGAAAAAAGTTGCAAAAATCATCAACA	2700
Qy	2701	AGAGATTTAGGTTAACTGTAATGTTTAAAGACATTTTATAGTGAAGAAATGTAGTA	2760
Dp	2701	AGAGATTTAGGTTAACTGTAATGTTTAAAGACATTTTATAGTGAAGAAATGTAGTA	2760
Qy	2761	TTTTCAGAGATTTGATTAACCTTATGCTCTTTTCAGGATCTTTCAAGAAAATGCTTTTG	2820
Dp	2761	TTTTCAGAGATTTGATTAACCTTATGCTCTTTTCAGGATCTTTCAAGAAAATGCTTTTG	2820
Qy	2821	GGGCTACAGAAAGCTTTAGAAAAATTTGAAGTGAATAAGGCAAAATAAGAAAAAT	2880
Dp	2821	GGGCTACAGAAAGCTTTAGAAAAATTTGAAGTGAATAAGGCAAAATAAGAAAAAT	2880
Qy	2881	GGTTTTCACGAGCACTGATCTTTTACTTTGATTAATTTTATCTGCTCTTTCTTTT	2940
Dp	2881	GGTTTTCACGAGCACTGATCTTTTACTTTGATTAATTTTATCTGCTCTTTCTTTT	2940
Qy	2941	CTCTAGCTTAACAGACTTAAAGAAAGCATCATGTTCCATGATTTTCAACCCAGGGGA	3000
Dp	2941	CTCTAGCTTAACAGACTTAAAGAAAGCATCATGTTCCATGATTTTCAACCCAGGGGA	3000
Qy	3001	CTTCAAGCTGCCCCGTCGCCAGAAATTTCAAGAGTAAAGGCCAGAGTAAACATGATTTGT	3060
Dp	3001	CTTCAAGCTGCCCCGTCGCCAGAAATTTCAAGAGTAAAGGCCAGAGTAAACATGATTTGT	3060
Qy	3061	TTTTCAGGAAGAGATCTGAGGCAAGGTTTTCACTGGGTGAGTGAAGCAAGCTGATGTGATC	3120
Dp	3061	TTTTCAGGAAGAGATCTGAGGCAAGGTTTTCACTGGGTGAGTGAAGCAAGCTGATGTGATC	3120
Qy	3121	AAAGAAATTTAATGTGACTTTGTCTACAGAGGCCGCTTGTCTTCAAGGGCAATTACT	3180
Dp	3121	AAAGAAATTTAATGTGACTTTGTCTACAGAGGCCGCTTGTCTTCAAGGGCAATTACT	3180
Qy	3181	GAGGAGAGCTTCCCAAGTCTGCTCTGAGCAATGCTGTCTAATTTCCCTGGGAAAAAAAGT	3240
Dp	3181	GAGGAGAGCTTCCCAAGTCTGCTCTGAGCAATGCTGTCTAATTTCCCTGGGAAAAAAAGT	3240
Qy	3241	CAACACTTAAAAAAGTGTCTTTCTCTCTTCCCTTTCACCCGCTCTTTTCCCAATTC	3300
Dp	3241	CAACACTTAAAAAAGTGTCTTTCTCTCTTCCCTTTCACCCGCTCTTTTCCCAATTC	3300
Qy	3301	CCTAGACACAGAGAAAGCCCTCTTTTGGGGCAGCTCATCTTAATTGAACTTGGAGAA	3360
Dp	3301	CCTAGACACAGAGAAAGCCCTCTTTTGGGGCAGCTCATCTTAATTGAACTTGGAGAA	3360
Qy	3361	GCTGGGTGAAGGCTCTTAATGCGACAGTTTACAAAGGGATTAGCAGGTGAGTGAACATAG	3420
Dp	3361	GCTGGGTGAAGGCTCTTAATGCGACAGTTTACAAAGGGATTAGCAGGTGAGTGAACATAG	3420
Qy	3421	CTGGAGAGACTTTAGAGATGAGATGAGTCCCGCCCCCAATTTCAATTATTAACCAAGGT	3480
Dp	3421	CTGGAGAGACTTTAGAGATGAGATGAGTCCCGCCCCCAATTTCAATTATTAACCAAGGT	3480
Qy	3481	GAGACATCATTAAGAGTTTATGACCTCAGGACTGTGCAAGACCAATGGCCGACAGGGA	3540
Dp	3481	GAGACATCATTAAGAGTTTATGACCTCAGGACTGTGCAAGACCAATGGCCGACAGGGA	3540

Qy	3541	GAGGACATGATTA	CTTAAACAGCCTTGAAGAAAACAAACCTGCCTGCCTAATTA	3600
Ds	3541	GAGGACATGATTA	CTTAAACAGCCTTGAAGAAAACAAACCTGCCTGCCTAATTA	3600
Qy	3501	AATAGCCCACTTAAT	GTATTATACGCTTTCCTTCCTTCGATTCAAATCAGAAATTC	3666
Ds	3501	AATAGCCCACTTAAT	GTATTATACGCTTTCCTTCCTTCGATTCAAATCAGAAATTC	3666
Qy	3661	AAGAAATTAACAT	CTCTACTGACCCAAAGAACAAATTAATCACTTCAGGCCCTGTG	3720
Ds	3661	AAGAAATTAACAT	CTCTACTGACCCAAAGAACAAATTAATCACTTCAGGCCCTGTG	3720
Qy	3721	GGAGGCA	CAGTGGTAAGGCTCTCAACAGTTTTTATATCCCTCCCTAATCAAAAT	3780
Ds	3721	GGAGGCA	CAGTGGTAAGGCTCTCAACAGTTTTTATATCCCTCCCTAATCAAAAT	3780
Qy	3781	GACAGATTTTTGT	ATATGGCAACCTGAAATTTGCTGCTTCAATCCCTCCAGCTGGCTTAT	3840
Ds	3781	GACAGATTTTTGT	ATATGGCAACCTGGAATTTGCTGCTTCAATCCCTCCAGCTGGCTTAT	3840
Qy	3841	AGAAAGAACTGA	AGTTGGTTCTGCAATTAATGTGTACATGCAAAAGATGATAAATCCTAG	3900
Ds	3841	AGAAAGAACTGA	AGTTGGTTCTGCAATTAATGTGTACATGCAAAAGATGATAAATCCTAG	3900
Qy	3901	ATTTTTTAATAT	TTTGCAAAATATACAAATATGTGTGSAATTAATAATCTGCTTATCCAA	3966
Ds	3901	ATTTTTTAATAT	TTTGCAAAATATACAAATATGTGTGSAATTAATAATCTGCTTATCCAA	3966
Qy	3961	AGCTAATGATCAAT	TTTGTGTAAACAAACCAACTTGTGTAAATATATGTATTAAGATCCATGA	4020
Ds	3961	AGCTAATGATCAAT	TTTGTGTAAACAAACCAACTTGTGTAAATATATGTATTAAGATCCATGA	4020
Qy	4021	ATTCCCTTTT	AGTCAAGTGGGAAGTTGATGCTGCTTTTTCCTTATGTTACTCA	4080
Ds	4021	ATTCCCTTTT	AGTCAAGTGGGAAGTTGATGCTGCTTTTTCCTTATGTTACTCA	4080
Qy	4081	ATAGAGAGAAAGTA	ATGCTCAATATGTGTAAATTTAATTTTAAATAATATAGCTGAT	4140
Ds	4081	ATAGAGAGAAAGTA	ATGCTCAATATGTGTAAATTTAATTTTAAATAATATAGCTGAT	4140
Qy	4141	CCGAGTCAGTGGT	GTAAACAACTACTGTATCAACACAGTTACAGATTTCTTGTGCTCT	4200
Ds	4141	CCGAGTCAGTGGT	GTAAACAACTACTGTATCAACACAGTTACAGATTTCTTGTGCTCT	4200
Qy	4201	TCTCCACTCCCACT	GTCTTCACTPACTGSCCAAAAAGAAAAAATTTTATATAA	4266
Ds	4201	TCTCCACTCCCACT	GTCTTCACTPACTGSCCAAAAAGAAAAAATTTTATATAA	4266
Qy	4261	CTACTCA	CAAGCTAATATTTATTTATTTATTTATGCTGTTATTTATTTT	4320
Ds	4261	CTACTCA	CAAGCTAATATTTATTTATTTATTTATGCTGTTATTTATTTT	4320
Qy	4321	ACTGTAAACAGGAT	TTGAGGGGACATACAGTTTATTTTATTTATTTATTTAT	4380
Ds	4321	ACTGTAAACAGGAT	TTGAGGGGACATACAGTTTATTTTATTTTATTTATTTAT	4380
Qy	4381	TTATTTATTTAT	TTTGGAAATGGAATCTCTGACCCACGCTGAGTCAAGTGGCATC	4440
Ds	4381	TTATTTATTTAT	TTTGGAAATGGAATCTCTGACCCACGCTGAGTCAAGTGGCATC	4440
Qy	4441	TCAGATGACTGCA	ACCTCTGCTCTGAGTTCAAGCAACCTCTCTGCCCCCTTTTA	4500
Ds	4441	TCAGATGACTGCA	ACCTCTGCTCTGAGTTCAAGCAACCTCTCTGCCCCCTTTTA	4500
Qy	4501	TACTTCTTAAT	CTGTTTAGTCAATGTATCACTTATTTTTCAAATGCTGAGAAACATC	4560
Ds	4501	TACTTCTTAAT	CTGTTTAGTCAATGTATCACTTATTTTTCAAATGCTGAGAAACATC	4560
Qy	4561	TGCAATTAAGACA	CACTTTTATTTTATTTTAAAGTTCCTCATATCAATTTGGCATGAT	4620
Ds	4561	TGCAATTAAGACA	CACTTTTATTTTATTTTAAAGTTCCTCATATCAATTTGGCATGAT	4620
Qy	4621	AATCTTTTCAAG	TGGCTCGAAGGGGCAACCTTGGAACATATCTTGGAATCATGGCA	4680

Db 4621 AACTGTTTCAAGTGCGTCGGAACGGGGGCAACCCGGAACATACCTTGATACATGGGCA 4680
Qy 4681 CGATGGACACTCTGATCTCTCTCTGAGTCTGACCTTGATGTTGTGCAAGACCTT 4740
Db 4681 CCAATGGACACTCTGATCTCTCTCTGAGTCTGACCTTGATGTTGTGCAAGACCTT 4740
Qy 4741 TCCAGCCGCAAGTTTACACAGAAATCATTATCTTTCTTAGTACTTATGTTTCT 4800
Db 4741 TCCAGCCGCAAGTTTACACAGAAATCATTATCTTTCTTAGTACTTATGTTTCT 4800
Qy 4801 TTTTCATTTAATCTCTTCTGATGTTGAAATGTCGCCATGATTAATATATAG 4920
Db 4801 TTTTCATTTAATCTCTTCTGATGTTGAAATGTCGCCATGATTAATATATAG 4920
Qy 4861 TCCATTTAATGATTTCTCTCTGATGTTGAAATGTCGCCATGATTAATATAG 4920
Db 4861 TCCATTTAATGATTTCTCTCTGATGTTGAAATGTCGCCATGATTAATATAG 4920
Qy 4921 ATCTCAAGATACTGAAATCTTTCTGTTCTAATCTTTTAAATCATGTTTCTTAAT 4980
Db 4921 ATCTCAAGATACTGAAATCTTTCTGTTCTAATCTTTTAAATCATGTTTCTTAAT 4980
Qy 4981 CTATCTTTTCTTATATTTGT 5000
Db 4981 CTATCTTTTCTTATATTTGT 5000

RESULT 2
AX571875 5332 bp DNA linear PAT 29-NOV-2002
LOCUS AX571875 Sequence 3 from Patent WO02061060.
DEFINITION AX571875
ACCESSION AX571875
VERSION AX571875.1 GI:26004000
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Yan, C., Kechum, K., di Francesco, V. and Beasley, E.M.
Isolated human kinase proteins, nucleic acid molecules encoding
human kinase proteins, and uses thereof
Patent: WO 02061060-A 3 08-AUG-2002;
PE Corporation (NY) (US)
FEATURES
source 1..5332
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 5000; DB 6; Length 5332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 241 CCTGTGTCTATATAAATAATATAGCTGTGTCTTGACACAGGCTCAGCTAGCTACT 300
Qy 301 CGAAGACTGAGTGGAGATCACTGAGCCAGAGGCTGAGGCTGAGTGAACAGTGA 360
Db 301 CGAAGACTGAGTGGAGATCACTGAGCCAGAGGCTGAGGCTGAGTGAACAGTGA 360
Qy 361 TCACCAGCTGATTCACGCTCGAAGACAGAGGAGACCTGTGTTCCAAAAA 420
Db 361 TCACCAGCTGATTCACGCTCGAAGACAGAGGAGACCTGTGTTCCAAAAA 420
Qy 421 AAAAAAATGCAAGAAAGACATCATTAACCTTGACCTGGACATTAATTTATGATG 480
Db 421 AAAAAAATGCAAGAAAGACATCATTAACCTTGACCTGGACATTAATTTATGATG 480
Qy 481 AAATTCACATCTTTAGAGAAATAGCATTTCTGATTAATATATATATATAT 540
Db 481 AAATTCACATCTTTAGAGAAATAGCATTTCTGATTAATATATATATATAT 540
Qy 541 ATTATTAATTCGAATGGAATTAATATTTCTGAGAAACCTAGCTTCTCAGTTGT 600
Db 541 ATTATTAATTCGAATGGAATTAATATTTCTGAGAAACCTAGCTTCTCAGTTGT 600
Qy 601 CAGTCAAAACCTTTAATGCTTTGCGGAGTGGCTCAGGCTGTAAATCCAGCAC 660
Db 601 CAGTCAAAACCTTTAATGCTTTGCGGAGTGGCTCAGGCTGTAAATCCAGCAC 660
Qy 661 TTTGGAGAGCCGAGGCGGATGATCAGAAGTTAGAGATGAGACATCTGCTAACA 720
Db 661 TTTGGAGAGCCGAGGCGGATGATCAGAAGTTAGAGATGAGACATCTGCTAACA 720
Qy 721 CGGTAAACCTGTCTCTAATTAATTAACAAAAATTTGCGGAGTGGCTCAGAGCC 780
Db 721 CGGTAAACCTGTCTCTAATTAATTAACAAAAATTTGCGGAGTGGCTCAGAGCC 780
Qy 781 TGTAGTCCAGTGTCTCAGAGGCTGAGAGAGATGATGTAACCCGAGAGCGAG 840
Db 781 TGTAGTCCAGTGTCTCAGAGGCTGAGAGAGATGATGTAACCCGAGAGCGAG 840
Qy 841 CTTCAGATGAGCGGAGATTTGGCCCATCTGACCTCGGCGGACAGTGGAGACTCG 900
Db 841 CTTCAGATGAGCGGAGATTTGGCCCATCTGACCTCGGCGGACAGTGGAGACTCG 900
Qy 901 TCTCAAAAAAATTTTAAAAAATGTAATGCTTTGAGGCAAGTGTCTTCTTTTC 960
Db 901 TCTCAAAAAAATTTTAAAAAATGTAATGCTTTGAGGCAAGTGTCTTCTTTTC 960
Qy 961 TCTTCTTCTTTTCTTTTCTTTTCAAAAAATATCTAGATTGAATCTTGAATTGGCT 1020
Db 961 TCTTCTTCTTTTCTTTTCTTTTCAAAAAATATCTAGATTGAATCTTGAATTGGCT 1020
Qy 1021 TAAGTCTCTTCTTTGCGCAATTTTGAATGAAAAATATCATGCTCATTAATTAAT 1080
Db 1021 TAAGTCTCTTCTTTGCGCAATTTTGAATGAAAAATATCATGCTCATTAATTAAT 1080
Qy 1081 CTGAACATTTTAAAAAATCATATAGAGTTTCAAAATATCAAAATATCAATTAATGTTGT 1140
Db 1081 CTGAACATTTTAAAAAATCATATAGAGTTTCAAAATATCAAAATATCAATTAATGTTGT 1140
Qy 1141 GATTAATAGACATTAATCTTATTTTCCCTTAATATATATGTTTATATATCTTCATTC 1200
Db 1141 GATTAATAGACATTAATCTTATTTTCCCTTAATATATATGTTTATATATCTTCATTC 1200
Qy 1201 TGTCTCATCTTATGATTAATTAATTAATGAGCAATCTTGAATTAACAAGAAA 1260
Db 1201 TGTCTCATCTTATGATTAATTAATTAATGAGCAATCTTGAATTAACAAGAAA 1260
Qy 1261 AGGTGTCATTTGAAGACATCAAGCTGCAACCAATTTTAAGCAGATTCTTGACATGTT 1320
Db 1261 AGGTGTCATTTGAAGACATCAAGCTGCAACCAATTTTAAGCAGATTCTTGACATGTT 1320
Qy 1321 GTGCTGTTAATATAGTACTTATCATATAGTAAACATTAATTAATATATGGAAGAA 1380
Db 1321 GTGCTGTTAATATAGTACTTATCATATAGTAAACATTAATTAATATATGGAAGAA 1380

Db	1321	GTGCTGTTAAATATATGATCTTAAACAATATGTAACAACTTAAATTAACAATATATGTGGAAGAG	1380
Qy	1381	CAAGCAAGAAAGAAAGAAAGATTTTCATTTCAAACTCTCTCTCTCCATCACCAATTGGCTA	1440
Db	1381	CAAGCAAGAAAGAAAGAAAGATTTTCATTTCAAACTCTCTCTCTCCATCACCAATTGGCTA	1440
Qy	1441	ATATCATCATTTTGTATGACGTTAAGAACAAATATAGGCGCTCACCAATGTTTTTGAATTA	1500
Db	1441	ATATCATCATTTTGTATGACGTTAAGAACAAATATAGGCGCTCACCAATGTTTTTGAATTA	1500
Qy	1501	TGAATGAAATGCAACCCCTTCTAAGACTTTTGGATACACTATTTGTTTGAAGCAAGAGAT	1560
Db	1501	TGAATGAAATGCAACCCCTTCTAAGACTTTTGGATACACTATTTGTTTGAAGCAAGAGAT	1560
Qy	1561	GCAGTAGATATTTTCAACTTTTTTCTGTTTTTATGATTTCTGTGGTTTCTTTGACTACTAA	1620
Db	1561	GCAGTAGATATTTTCAACTTTTTTCTGTTTTTATGATTTCTGTGGTTTCTTTGACTACTAA	1620
Qy	1621	AAGTTACTAGGTATGCAAAATTTGTTTTTAAAGTCGAAAAACAAAATGCTTTCAAATATAA	1680
Db	1621	AAGTTACTAGGTATGCAAAATTTGTTTTTAAAGTCGAAAAACAAAATGCTTTCAAATATAA	1680
Qy	1681	GGTAGGGAGAAAAATATCCTCTCAACATGTCACATTTTAGCACAGAAAAACCTAATATCAA	1740
Db	1681	GGTAGGGAGAAAAATATCCTCTCAACATGTCACATTTTAGCACAGAAAAACCTAATATCAA	1740
Qy	1741	TATCACCATCATGATATCATATTAATATCATTTGCATATGATTAAGCAATGTCAATCCCTAA	1800
Db	1741	TATCACCATCATGATATCATATTAATATCATTTGCATATGATTAAGCAATGTCAATCCCTAA	1800
Qy	1801	AAACTATATGATACCAATATGCACTAACTTTGTGGCCAGAACAAAGACCTTAACTGTGCCAA	1860
Db	1801	AAACTATATGATACCAATATGCACTAACTTTGTGGCCAGAACAAAGACCTTAACTGTGCCAA	1860
Qy	1861	TTTTATTTCTATTCANATTAACAGCTGCTCGTTTTGAGTTGTGACATATGCAAGCAA	1920
Db	1861	TTTTATTTCTATTCANATTAACAGCTGCTCGTTTTGAGTTGTGACATATGCAAGCAA	1920
Qy	1921	TCCCTGTCTGATGTGAGATTTCTTGCACTGATTAAGAAAACTGCTGAAGTTGTGAGGCT	1980
Db	1921	TCCCTGTCTGATGTGAGATTTCTTGCACTGATTAAGAAAACTGCTGAAGTTGTGAGGCT	1980
Qy	1981	GCTTCAGGACAGCATCATGTGATCATATGAAGCTCCACGCTGCTGAACCTCTGGCAA	2040
Db	1981	GCTTCAGGACAGCATCATGTGATCATATGAAGCTCCACGCTGCTGAACCTCTGGCAA	2040
Qy	2041	AAAGGGAGAGAACAAAGATAGGAGAGGACAGTGGGGGAAAGGTTCAAGTGGGGTTTTCTC	2100
Db	2041	AAAGGGAGAGAACAAAGATAGGAGAGGACAGTGGGGGAAAGGTTCAAGTGGGGTTTTCTC	2100
Qy	2101	CTTGAAACCTACAAAGATTTATGGGTCAAGAGCTGTGTGCAAAAGACTGTACAGCTTGATGCA	2160
Db	2101	CTTGAAACCTACAAAGATTTATGGGTCAAGAGCTGTGTGCAAAAGACTGTACAGCTTGATGCA	2160
Qy	2161	GCTGCTAACCATTTGTTCAAGGGAGGCGAGGACACAGCTGTGGAAGAGTCAAGCTTGAGA	2220
Db	2161	GCTGCTAACCATTTGTTCAAGGGAGGCGAGGACACAGCTGTGGAAGAGTCAAGCTTGAGA	2220
Qy	2221	CCAGGGAGCGTGGCTTCAAGGATTTTGTATCCCAAGAGAGAGCATCTTTCTCTATTTGATA	2280
Db	2221	CCAGGGAGCGTGGCTTCAAGGATTTTGTATCCCAAGAGAGAGCATCTTTCTCTATTTGATA	2280
Qy	2281	AACCAAGAGATTCAAGACACTCCCTTTTTTGTAGCGGAGTCTAATCTTCTGCGGTAAGGCT	2340
Db	2281	AACCAAGAGATTCAAGACACTCCCTTTTTTGTAGCGGAGTCTAATCTTCTGCGGTAAGGCT	2340
Qy	2341	AAACCAATTAATAATGAAAAATTTCTATTTAAAGTCACAGAAAAATTAATGCGCTGTATATCAA	2400
Db	2341	AAACCAATTAATAATGAAAAATTTCTATTTAAAGTCACAGAAAAATTAATGCGCTGTATATCAA	2400
Qy	2401	TTTGGGGAAATTTCTGTATTAACAAAGGAAAAATTAATCTTGTGGCTTTGGGCTGACGAA	2460
Db	2401	TTTGGGGAAATTTCTGTATTAACAAAGGAAAAATTAATCTTGTGGCTTTGGGCTGACGAA	2460

Qy	2461	ACTACCTGGCTTGAAGTCGAGAAAGTAGTCTCTCAAAATCTCTAAAGTCTCTAAATTAC	2520
Dp	2461	ACTACCTGGCTTGAAGTCGAGAAAGTAGTCTCTCAAAATCTCTAAAGTCTCTAAATTAC	2520
Qy	2521	AGAGCTTAAACTTTAAAGGCAAGCTGCAGTATTAGTTGTAATGCTATGATTTTGAAACTT	2580
Dp	2521	AGAGCTTAAACTTTAAAGGCAAGCTGCAGTATTAGTTGTAATGCTATGATTTTGAAACTT	2580
Qy	2581	TAGTAATTTAGTCAATGATTTATAGCAATGCCATGATTTATTCCTCTACAGAAATTAATTA	2640
Dp	2581	TAGTAATTTAGTCAATGATTTATAGCAATGCCATGATTTATTCCTCTACAGAAATTAATTA	2640
Qy	2641	AGTGACATGAAAAAAGGCAAGCTTCAACAGAAAAAAGTTGCAAAACATCATCAAA	2700
Dp	2641	AGTGACATGAAAAAAGGCAAGCTTCAACAGAAAAAAGTTGCAAAACATCATCAAA	2700
Qy	2701	AGAGATTTAGGTTAACTGTAATGTTAAAGACACTTTTTAGTGTAAGAAAGATGTAGTA	2760
Dp	2701	AGAGATTTAGGTTAACTGTAATGTTAAAGACACTTTTTAGTGTAAGAAAGATGTAGTA	2760
Qy	2761	TTTCAAGAGTTGAATACCATATGCTCTTTTCAAGGATCTTTCAAGAAAAAGTGCCTTTTG	2820
Dp	2761	TTTCAAGAGTTGAATACCATATGCTCTTTTCAAGGATCTTTCAAGAAAAAGTGCCTTTTG	2820
Qy	2821	GGGGTACAGGAAGCTTAGAAAACATTTTGAAGTGAATAATGAGGCAATATAAGAAAAAT	2880
Dp	2821	GGGGTACAGGAAGCTTAGAAAACATTTTGAAGTGAATAATGAGGCAATATAAGAAAAAT	2880
Qy	2881	GGTTTTCACAGGCATGCAATCTTTTCACTTGCATPAAATTTATTTTCCTCTCTTTCTTTT	2940
Dp	2881	GGTTTTCACAGGCATGCAATCTTTTCACTTGCATPAAATTTATTTTCCTCTCTTTCTTTT	2940
Qy	2941	CTTAGCTTAAACAGACTTAAAGAAAGCATCATGTGTCATGACCTTGATTTCAACCCAGGGGA	3000
Dp	2941	CTTAGCTTAAACAGACTTAAAGAAAGCATCATGTGTCATGACCTTGATTTCAACCCAGGGGA	3000
Qy	3001	CTTCAAGCTGCCGCTGCCAGAAAGTTCAAGATPAAAGGCCAGGATTAACAGTGAATGT	3060
Dp	3001	CTTCAAGCTGCCGCTGCCAGAAAGTTCAAGATPAAAGGCCAGGATTAACAGTGAATGT	3060
Qy	3061	TTTTCAGGAAGAGATCTGAAGGACGAGGTTTTCAGTGGGTGAGTGAAGCAGCTGATTTGATC	3120
Dp	3061	TTTTCAGGAAGAGATCTGAAGGACGAGGTTTTCAGTGGGTGAGTGAAGCAGCTGATTTGATC	3120
Qy	3121	AAGAAAGATTTAATGTAGCTTGTCTAACGAGGCGGCGCTTGCTTCACAGGCAATTAATCT	3180
Dp	3121	AAGAAAGATTTAATGTAGCTTGTCTAACGAGGCGGCGCTTGCTTCACAGGCAATTAATCT	3180
Qy	3181	GAGGAGACCTTCCCAATCTGCTCTGGCAATGCTGTCTAATTTTCCCTGGGGAAAAAAGT	3240
Dp	3181	GAGGAGACCTTCCCAATCTGCTCTGGCAATGCTGTCTAATTTTCCCTGGGGAAAAAAGT	3240
Qy	3241	CAACACATPAAAAAAGTGTCTTTCTCTCTTCCCTTCAACCGGCTCCTTTTCCCAATTC	3300
Dp	3241	CAACACATPAAAAAAGTGTCTTTCTCTCTTCCCTTCAACCGGCTCCTTTTCCCAATTC	3300
Qy	3301	CCTTAGACCAAGGAAGGCTCCCTCTTTTGGGGACGCTCATCTTAATTGAACCTTGAGAA	3360
Dp	3301	CCTTAGACCAAGGAAGGCTCCCTCTTTTGGGGACGCTCATCTTAATTGAACCTTGAGAA	3360
Qy	3361	GCTGGGTGAAGGCTCTTAATGCAAGGTTTCAAGGGGATTAGCAGGTGAGTGAACAATAG	3420
Dp	3361	GCTGGGTGAAGGCTCTTAATGCAAGGTTTCAAGGGGATTAGCAGGTGAGTGAACAATAG	3420
Qy	3421	CTGGGAGAGACTTTAGAGATGAGAGTCCGCGCCCCCAATTTCAATTAATPAAAGCAGGT	3480
Dp	3421	CTGGGAGAGACTTTAGAGATGAGAGTCCGCGCCCCCAATTTCAATTAATPAAAGCAGGT	3480
Qy	3481	GAGACATCATAGAAAGTTCAATAGCACTGAGACCTGTGCAAGAACCAATGCGCCAGAGGA	3540
Dp	3481	GAGACATCATAGAAAGTTCAATAGCACTGAGACCTGTGCAAGAACCAATGCGCCAGAGGA	3540

QY 3541 GAGAGACATGATTAATTAAACAGCCTTGAAGAAAAACAACCTGCTGCTCTTAATTAA 3600
 DB 3541 GAGAGACATGATTAATTAAACAGCCTTGAAGAAAAACAACCTGCTGCTCTTAATTAA 3600
 QY 3601 AATCAGCCCACTTAATGTTTATCAGCCTTCCCTCTGCAATCAATCAGAGATTCA 3660
 DB 3601 AATCAGCCCACTTAATGTTTATCAGCCTTCCCTCTGCAATCAATCAGAGATTCA 3660
 QY 3661 AATCAGCCCACTTAATGTTTATCAGCCTTCCCTCTGCAATCAATCAGAGATTCA 3660
 DB 3661 AATCAGCCCACTTAATGTTTATCAGCCTTCCCTCTGCAATCAATCAGAGATTCA 3660
 QY 3661 AAGAAATAGACATCTCTACTACTGACCCAAAGAACAAATTAATCACTCTTCAGGCTGTG 3720
 DB 3661 AAGAAATAGACATCTCTACTACTGACCCAAAGAACAAATTAATCACTCTTCAGGCTGTG 3720
 QY 3721 GAGAGCAAGTTGGTAAAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAAT 3780
 DB 3721 GAGAGCAAGTTGGTAAAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAAT 3780
 QY 3781 GAGAGCAAGTTGGTAAAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAAT 3840
 DB 3781 GAGAGCAAGTTGGTAAAGCGTCTTAACAGGTTTTTATATCCCTCCCTTAATCAAT 3840
 QY 3841 AGAAGAACTGAAGTTGGTCTGCAAAATTATGTATCATGCAAAAGATGATTAATCTAG 3900
 DB 3841 AGAAGAACTGAAGTTGGTCTGCAAAATTATGTATCATGCAAAAGATGATTAATCTAG 3900
 QY 3901 ATTTTATATTTTGCAGAAATACCAAAATGTCTGAGAAATTAATAATCTTATCCAAA 3960
 DB 3901 ATTTTATATTTTGCAGAAATACCAAAATGTCTGAGAAATTAATAATCTTATCCAAA 3960
 QY 3961 AGCTAATGATTAATTTTGTAAACAACCACTTGTAAATATATGTAATTAATAGCATGA 4020
 DB 3961 AGCTAATGATTAATTTTGTAAACAACCACTTGTAAATATATGTAATTAATAGCATGA 4020
 QY 4021 ATTCCTCTTTTGTCAAGGTGGGAAAGTTGATGCTGCTTTTCTTTATGTTACTCCA 4080
 DB 4021 ATTCCTCTTTTGTCAAGGTGGGAAAGTTGATGCTGCTTTTCTTTATGTTACTCCA 4080
 QY 4081 ATGAGAGAAAGTAATGCTCAATGCTGTAATTTTAAATTTTAAATATAGCGAT 4140
 DB 4081 ATGAGAGAAAGTAATGCTCAATGCTGTAATTTTAAATTTTAAATATAGCGAT 4140
 QY 4141 CCGAGTGCAGTGGTGTAAACAACCACTTGTATCAACAACAGTTTCTTTGTTCT 4200
 DB 4141 CCGAGTGCAGTGGTGTAAACAACCACTTGTATCAACAACAGTTTCTTTGTTCT 4200
 QY 4201 TCTTCACCTCCACTGCTTCACTTACTGCGCCAAAAAGAAAAATTTTATTA 4260
 DB 4201 TCTTCACCTCCACTGCTTCACTTACTGCGCCAAAAAGAAAAATTTTATTA 4260
 QY 4261 CTACTACAAGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4320
 DB 4261 CTACTACAAGCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4320
 QY 4321 ACTTGTAAACAGATTTGAGGGAGATACAGTTTATTTTATTTATTTATTAAT 4380
 DB 4321 ACTTGTAAACAGATTTGAGGGAGATACAGTTTATTTTATTTATTTATTAATTAAT 4380
 QY 4381 TTATTTATTTTATTTGAGATGATCTCTGTCAACCCAGCTGAGAGTGCAGATC 4440
 DB 4381 TTATTTATTTTATTTGAGATGATCTCTGTCAACCCAGCTGAGAGTGCAGATC 4440
 QY 4441 TCAATGACCTGCACTCTGCTCTGCTGAGTCAAGCACTCTCTGCGCCCTGAGCTTGA 4500
 DB 4441 TCAATGACCTGCACTCTGCTCTGCTGAGTCAAGCACTCTCTGCGCCCTGAGCTTGA 4500
 QY 4501 TACTTTCTTAATCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
 DB 4501 TACTTTCTTAATCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
 QY 4561 TGCATTAAGAGCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4620
 DB 4561 TGCATTAAGAGCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 4620
 QY 4621 AACTGTTTCAAGGTGCTCGGAACGGGGGCACTCTGGAACATCTTGATATCATGAGCA 4680

DB 4621 AACTGTTTCAAGGTGCTCGGAACGGGGGCACTCTGGAACATCTTGATATCATGAGCA 4680
 QY 4681 CCATGACACTTCTGATCTCTCTCTGAGTCTGACTTGAATGTTCTGCAAGACTT 4740
 DB 4681 CCATGACACTTCTGATCTCTCTCTGAGTCTGACTTGAATGTTCTGCAAGACTT 4740
 QY 4741 TCCAGCCGGAAGTTTACACAAATTCATCTTTCTTCTAGTACTTATGTTTCT 4800
 DB 4741 TCCAGCCGGAAGTTTACACAAATTCATCTTTCTTCTAGTACTTATGTTTCT 4800
 QY 4801 TTTTCAATTAATCTCTTCTATCTAGTGGAAATTTATTTATTTATTTATTTATTTATTTAT 4860
 DB 4801 TTTTCAATTAATCTCTTCTATCTAGTGGAAATTTATTTATTTATTTATTTATTTATTTAT 4860
 QY 4861 TCCATTTATAGATTTCTTTTCTCTGATGTTGAAATGCTGCAATGATTAATTAAT 4920
 DB 4861 TCCATTTATAGATTTCTTTTCTCTGATGTTGAAATGCTGCAATGATTAATTAAT 4920
 QY 4921 ATCTCAGAAATCTGAAATCTTCTGCTGATGTTGAAATGCTGCAATGATTAATTAAT 4980
 DB 4921 ATCTCAGAAATCTGAAATCTTCTGCTGATGTTGAAATGCTGCAATGATTAATTAAT 4980
 QY 4981 CTATCTTTCTTATTTTGT 5000
 DB 4981 CTATCTTTCTTATTTTGT 5000

RESULT 3
 AC007242/c 112748 bp DNA linear PRI 30-SEP-2000
 LOCUS Homo sapiens BAC clone RP11-182H9 from 2, complete sequence.
 DEFINITION AC007242.3 GI:5931465
 ACCESSION AC007242.3 GI:5931465
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 112748)
 Suleston,J.E. and Waterston,R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE
 99063792
 PUBMED
 9847074
 REFERENCE
 2 (bases 1 to 112748)
 Kozlowicz,A., Stoneking,T., Hawkins,M. and Hawrysko,C.
 The sequence of Homo sapiens BAC clone RP11-182H9
 Unpublished
 JOURNAL
 3 (bases 1 to 112748)
 Waterston,R.H.
 Direct Submision
 REFERENCE
 Submitted (07-APR-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 112748)
 Waterston,R.H.
 Direct Submision
 REFERENCE
 Submitted (28-SEP-1999) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 112748)
 Waterston,R.
 Direct Submision
 REFERENCE
 Submitted (30-SEP-2000) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Sep 28, 1999 this sequence version replaced gi:5001521.
 COMMENT
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sepiens@wustl.wustl.edu
 ----- Summary Statistics

Center project name: H_NH0182H09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moom, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-53H10, 200 bp overlap; the clone sequenced to the right is RP11-309N8. Actual start of this clone is at base position 100126 of RP11-53H10; actual end is at 112748 of RP11-182H9.

The reads used to determine the sequence at the region of base positions 31840 to 112748 consists of numerous polymorphic base changes represented from both NH0182H09 and RP11-309N8. It is unknown which base calls belong to which clone.

Location/Qualifiers

1. 112748
/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

/clone="RP11-182H9"

/clone_11b="RPCI-11"

1033. 1331
/rpt_family="Alu"

3574. 3737
/rpt_family="MIR"

3814. 4009
/rpt_family="MIR"

4048. 4358
/rpt_family="MIR"

6038. 6457
/rpt_family="Alu"

6477. 6680
/rpt_family="MIR"

6681. 6978
/rpt_family="L2"

7386. 7542
/rpt_family="Alu"

7543. 7567
/rpt_family="Alu"

7568. 7702
/rpt_family="Alu"

8575. 8596
/rpt_family="AT-rich"

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

8599. 9529
/rpt_family="L1"

9685. 9875
/rpt_family="MIR"

10125. 10238
/rpt_family="MIR"

10357. 10458
/rpt_family="MIR"

10484. 10784
/rpt_family="MIR"

10898. 11020
/rpt_family="MIR"

11835. 12166
/db_xref="GI:1234078"

12082. 12118
/rpt_family="CA)n"

12210. 12521
/rpt_family="Alu"

13068. 13394
/rpt_family="L2"

13429. 13479
/rpt_family="Alu"

13511. 13812
/rpt_family="Alu"

13967. 14092
/rpt_family="L2"

14111. 14139
/rpt_family="AT-rich"

14206. 14515
/rpt_family="Alu"

14516. 14559
/rpt_family="Alu"

14723. 14767
/rpt_family="AT-rich"

15360. 15575
/rpt_family="MIR"

15691. 15734
/rpt_family="L2"

16277. 16577
/rpt_family="Alu"

16608. 16649
/rpt_family="L2"

16793. 16804
/rpt_family="MIR"

16919. 17003
/rpt_family="MIR"

19659. 19795
/rpt_family="GA-rich"

20648. 20707
/rpt_family="L2"

20881. 21187
/rpt_family="Alu"

22002. 22304
/rpt_family="Alu"

23542. 23699
/rpt_family="L2"

23944. 24249
/rpt_family="Alu"

25966. 26170
/rpt_family="L1"

26171. 26461
/rpt_family="Alu"

26476. 26808
/rpt_family="MIR"

26809. 27119
/rpt_family="Alu"

27120. 27185
/rpt_family="MIR"

27186. 27530
/rpt_family="L1"

27858. 27971
/rpt_family="Alu"

28007. 28310

```
repeat_region /rpt family="Alu" 28872..29076
repeat_region /rpt family="MIR" 29419..29758
repeat_region /rpt family="L1" 30018..30265
repeat_region /rpt family="Alu" 30668..30739
misc_feature 30668..30739
repeat_region /note=match to ESR AA436054 (NID:ig2140968) zu01c12.r1
repeat_region 31066..31155
```

Query Match	99.5%	Score 4973.4;	DB 9;	Length 112748;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 4999;	Conservative	0;	Mismatches 1;	Indels 5;
				Gaps 2;

QY	1	TATAGGCCAATGCTGTGGCTCAACGGTGTATATCCAGACACTTTGGAGACAGAGAAATCG	60
Db	73582	TATAGGCCAATGCTGTGGCTCAACGGTGTATATCCAGACACTTTGGAGAGCGAGAGATCG	73523
QY	61	CTTGAGCTCAGGAATTGGAGACAAGCTTAACGTAAATCTGTCTGAACAA	120
Db	73522	CTTGAGCTCAGGAATTGGAGACAAGCTTAACGTAAATCTGTCTGTACAA	73463
QY	121	TAAATAAAAGATTTTCCAGGCATGTGTGGCGTCACTCCCAAGTCCAGCTATTTGGAGGC	180
Db	73462	TAAATAAAAGATTTTCCAGGCATGTGTGGCGTCACTCCCAAGTCCAGCTATTTGGAGGC	73403
QY	181	TGAGGTAGAGGAATGCTTGAAGCCGAGAGTTGAAAGACAAGCTTAGGGAACATGTGAGA	240
Db	73402	TGAGGTAGAGGAATGCTTGAAGCCGAGAGTTGAAAGACAAGCTTAGGGAACATGTGAGA	73343
QY	241	CCCTGTGCTATATAAAATATATAGCTGTGTGTCTTGGACAGGCTCGAGCTAGCTACT	300
Db	73342	CCCTGTGCTATATAAAATATATAGCTGTGTGTCTTGGACAGGCTCGAGCTAGCTACT	73283
QY	301	CGAAGACTAGAGTGGAGAGATCACTGAGGCCAGAGAGCTGAGGCTGCAGTGAACAGTGA	360
Db	73282	CGAAGACTAGAGTGGAGAGATCACTGAGGCCAGAGAGCTGAGGCTGCAGTGAACAGTGA	73223
QY	361	TCACCCAGCTGATTTCCAGCCTGGAAGACAGAGGAGACCTCTGTTTCCAAAAA	420
Db	73222	TCACCCAGCTGATTTCCAGCCTGGAAGACAGAGGAGACCTCTGTTTCCAAAAA	73163
QY	421	AAAAAATAGCAGAAAAAGACATCAATAACTTGACCTGGAGACATACTTTATGTATG	480
Db	73162	AAAAAATAGCAGAAAAAGACATCAATAACTTGACCTGGAGACATACTTTATGTATG	73103
QY	481	AAATTCACAATCTTTTAGAAGAAATTAGCAATTTCTGATTAATGTATTAATTAATTT	540
Db	73102	AAATTCACAATCTTTTAGAAGAAATTAGCAATTTCTGATTAATGTATTAATTAATTT	73043
QY	541	ATTATAAATTCAAATAGGAATTAATTAATTTCTGAGAACTAGCTCTCACTCTCTCAGTTGT	600
Db	73042	ATTATAAATTCAAATAGGAATTAATTAATTTCTGAGAACTAGCTCTCACTCTCTCAGTTGT	72983
QY	601	CAGTCAAAACTTTATAGTCTTTGGCCGGGTGCGGTGCTCAACGCTGTAAATCCAGGAC	660
Db	72982	CAGTCAAAACTTTATAGTCTTTGGCCGGGTGCGGTGCTCAACGCTGTAAATCCAGGAC	72923
QY	661	TTTGGAGAGCCGAGGCCGAGTGAATCACAAGGTTTAGAGATTCAGAACCATCTGGCTTACA	720
Db	72922	TTTGGAGAGCCGAGGCCGAGTGAATCACAAGGTTTAGAGATTCAGAACCATCTGGCTTACA	72863
QY	721	CGGTGAACCTGTGCTCTACTTAAATAACAAAAATTAGCCGGGTGCGGTGCCAGACGCC	780
Db	72862	CGGTGAACCTGTGCTCTACTTAAATAACAAAAATTAGCCGGGTGCGGTGCCAGACGCC	72803
QY	781	TGTATGCTCCAGTGTCTCAGAGAGCTGAGAGCGAGAAATGTGTGAACCCCGAGAGCGGAG	840
Db	72802	TGTATGCTCCAGTGTCTCAGAGAGCTGAGAGCGAGAAATGTGTGAACCCCGAGAGCGGAG	72743
QY	841	CTTGCAGTAGGCCGAGATTGGCCCACTGACTCCAGCTGGGCGACAGTGCAGACTCTG	900

Db	72742	CTTGCAGTGGCCGAGATGGCCCACTGCACCTCCAGCCTGGGCGACAGTGGGAGACTCTG	72683
Qy	901	TCGC -AAAAAAAAAAAAAAAAAGTTGAATGAGCTTTGAGCCCAAGTACGTCCTCTTT	959
Db	72682	TCGCAAAAAAAAAAAAAAAAAGTTGAAATGAGCTTTGAGCCCAAGTACGTCCTCTTT	72623
Qy	960	CTTCCTCTCTTTTTTTTTTTTTTTTTCAAAAAATCTCTGATTTGAATCTTGGAATTTGGC	1019
Db	72622	CTTCCTCTCTTTTTTTTTTTTTTTTTCAAAAAATCTCTGATTTGAATCTTGGAATTTGGC	72563
Qy	1020	TTAAGTCCTCTCTCTTGCGCAATTTTGAATAGAAAAAATACATGCTCATATTAATTA	1079
Db	72562	TTAAGTCCTCTCTCTTGCGCAATTTTGAATAGAAAAAATACATGCTCATATTAATTA	72503
Qy	1080	CCTGACATTTTAAAAAACCATCATGAGGTTCAAATATCTCAATTTCTCAATTTGTTG	1139
Db	72502	CCTGACATTTTAAAAAACCATCATGAGGTTCAAATATCTCAATTTCTCAATTTGTTG	72443
Qy	1140	TGATATATGACATTAACCTATTTTTTCCCTTAATATATGATTTGTTATATCTCCAT	1199
Db	72442	TGATATATGACATTAACCTATTTTTTCCCTTAATATATGATTTGTTATATCTCCAT	72383
Qy	1200	CTGCTCACTTTATGATTTAGTATATATATATAGTGGCAATTAATCTTAGAATCTTAACAGAA	1259
Db	72382	CTGCTCACTTTATGATTTAGTATATATATATAGTGGCAATTAATCTTAGAATCTTAACAGAA	72323
Qy	1260	AAGTGTGCATTTGAACATTAACGACGCGCAACCAATTAAGCCAGATTCCTTGACATG	1319
Db	72322	AAGTGTGCATTTGAACATTAACGACGCGCAACCAATTAAGCCAGATTCCTTGACATG	72263
Qy	1320	TGTGCTGTTAATATAGTACTTTTACATATAGTAAACATTAATTAATATATATGGAAGAA	1379
Db	72262	TGTGCTGTTAATATAGTACTTTTACATATAGTAAACATTAATTAATATATGGAAGAA	72203
Qy	1380	GCAACCAAGAAAGGAAGAAAGTATTTCAATCAAACTCCTCTCTGCATCACCATTGGCT	1439
Db	72202	GCAACCAAGAAAGGAAGAAAGTATTTCAATCAAACTCCTCTCTGCATCACCATTGGCT	72143
Qy	1440	AATATCATCATTTTGTACAGTTAAGAACAAACATAGTGTCTCACCACATAGTTTGAATTA	1499
Db	72142	AATATCATCATTTTGTACAGTTAAGAACAAACATAGTGTCTCACCACATAGTTTGAATTA	72083
Qy	1500	ATGAATGAATGGCAACCTTTTAAGACTATTTGGAATACATATGTTTGAAGGCAAGAGA	1559
Db	72082	ATGAATGAATGGCAACCTTTTAAGACTATTTGGAATACATATGTTTGAAGGCAAGAGA	72022
Qy	1560	TGCAGTAGATTTTCACTTTTTTCCGTATTATGATTTCTGTGGTTCTTTGACTATCA	1619
Db	72022	TGCAGTAGATTTTCACTTTTTTCCGTATTATGATTTCTGTGGTTCTTTGACTATCA	71962
Qy	1620	AAAGTTAGCTAGTAGCAAAATTTGTTTAAAGTCTGAAAAACAAATAGCTTTCAGATAAA	1679
Db	71962	AAAGTTAGCTAGTAGTAGCAAAATTTGTTTAAAGTCTGAAAAACAAATAGCTTTCAGATAAA	71902
Qy	1680	AGGTAGGGAAGAAATATCTCCCAACATGTCCAATTAGACACAGAAACCTTAATATCA	1739
Db	71902	AGGTAGGGAAGAAATATCTCCCAACATGTCCAATTAGACACAGAAACCTTAATATCA	71844
Qy	1740	ATATCAACATCATGATATCATATTAATATATCATTTGCATATGAATGAATGTCATTCCTTA	1799
Db	71842	ATATCAACATCATGATATCATATTAATATATCATTTGCATATGAATGAATGTCATTCCTTA	71782
Qy	1800	AAAACTATGTATATCAAAATGACATTAATCTTGCGCAGAAACAAGAACCTTAATCTGCGAA	1859
Db	71782	AAAACTATGTATATCAAAATGACATTAATCTTGCGCAGAAACAAGAACCTTAATCTGCGAA	71722
Qy	1860	ATTTTATTCATTCATAATACAGCTGCTCGTTTTCAGTTGTGCAATCTGAATGCAAGCA	1919
Db	71722	ATTTTATTCATTCATAATACAGCTGCTCGTTTTCAGTTGTGCAATCTGAATGCAAGCA	71662
Qy	1920	ATCCGTGTGATGTGGAAGTTTCTTGCACTGATTAAGAAAAACGTGTAAGTTGGAAGC	19792


```

QY 4140 TCCGAGTGCAGTGTGTTTACACTTATGATCACAACACGATTACAGATTCTTGTTC 4199
DB 69442 TCCGAGTGCAGTGTGTTTACACTTATGATCACAACACGATTACAGATTCTTGTTC 69383
QY 4200 TTCTCCACTCCCACTGCTCTCACTTAACTGGCCAAAAAGAAAAAATTTATATA 4259
DB 69382 TTCTCCACTCCCACTGCTCTCACTTAACTGGCCAAAAAGAAAAAATTTATATA 69323
QY 4260 ACTACTACAAAGACTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4319
DB 69322 ACTACTACAAAGACTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 69263
QY 4320 TACTGTGTTAAACAGAGATTGTAGGGGACATACAGTTTATTTATTTATTTATTA 4377
DB 69262 TACTGTGTTAAACAGAGATTGTAGGGGACATACAGTTTATTTATTTATTTATTTATTA 69203
QY 4378 --TATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTA 4435
DB 69202 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTA 69143
QY 4436 CGATCTCAGATGACTGCACACTCTGCTGCTGAGTTCAAGCAACTCTCTGCCCCCTGACC 4495
DB 69142 CGATCTCAGATGACTGCACACTCTGCTGCTGAGTTCAAGCAACTCTCTGCCCCCTGACC 69083
QY 4496 CTTTAACTCTTCTTAAATCTGTTTATGTCATGCTGTCACCTTAACTTTTTCATGCTGAGA 4555
DB 69082 CTTTAACTCTTCTTAAATCTGTTTATGTCATGCTGTCACCTTAACTTTTTCATGCTGAGA 69023
QY 4556 ACATCTGCATTAAGGACCACTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 4615
DB 69022 ACATCTGCATTAAGGACCACTTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 68963
QY 4616 ATGATTAATCTGTTTCAAGTGTGCTCGGAACGGGGGACCCCTGGAACATTACTTGATTAAT 4675
DB 68962 ATGATTAATCTGTTTCAAGTGTGCTCGGAACGGGGGACCCCTGGAACATTACTTGATTAAT 68903
QY 4676 GGGGACCATGAGACACTTCTGATCTCTCTCTGATGTTTCACTTTATTTATTTATTTATTTATTT 4735
DB 68902 GGGGACCATGAGACACTTCTGATCTCTCTCTGATGTTTCACTTTATTTATTTATTTATTTATTT 68843
QY 4736 ACCCTTCCAGCCCGAAGTTTACACAGAAATCACTTATCTTTCTGAGTTACTTTATGT 4795
DB 68842 ACCCTTCCAGCCCGAAGTTTACACAGAAATCACTTATCTTTCTGAGTTACTTTATGT 68783
QY 4796 TTTCTTTTCACTTAACTCTTCTTCACTGAGGAAATTTATTTATTTATTTATTTATTTATTTATTT 4855
DB 68782 TTTCTTTTCACTTAACTCTTCTTCACTGAGGAAATTTATTTATTTATTTATTTATTTATTTATTT 68723
QY 4856 CCAGCTCCATTTATTAATTTCTTTCTCTGAGTGTGAAGTGTGAAGTGTGAGTATTAAT 4915
DB 68722 CCAGCTCCATTTATTAATTTCTTTCTCTGAGTGTGAAGTGTGAGTATTAAT 68663
QY 4916 ATTAGATCTCAGAAATCTGAAATCTTCTGTTCTTAATCTTTTAAAAATCATGTTTCC 4975
DB 68662 ATTAGATCTCAGAAATCTGAAATCTTCTGTTCTTAATCTTTTAAAAATCATGTTTCC 68603
QY 4976 TTAATCTAATCTTTTCTTATTAATTTGT 5000
DB 68602 TTAATCTAATCTTTTCTTATTAATTTGT 68578

```

```

RESULT 4
AC116995 173131 bp DNA linear ROD 05-NOV-2003
LOCUS Mus musculus BAC clone RP23-3C16 from 1, complete sequence.
AC116995
AC116995 GI:21844700
VERSION HTG
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 173131)

```

```

AUTHORS Holmes, A., Haglund, K. and Spalding, L.
TITLE The sequence of Mus musculus BAC clone RP23-3C16
JOURNAL Unpublished (2001)
REFERENCE 2 (bases 1 to 173131)
AUTHORS Wilson, R.
TITLE Sequencing of Mus musculus
JOURNAL Unpublished (2001)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) Genome Sequencing Center, 4444 Forest Park
4 (bases 1 to 173131)
McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2002) Genome Sequencing Center, 4444 Forest Park
5 (bases 1 to 173131)
McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2002) Genome Sequencing Center, 4444 Forest Park
6 (bases 1 to 173131)
McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 16, 2002 this sequence version replaced gi:20389753.
COMMENT
-----
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.wustl.edu
-----
Summary Statistics
Center project name: M_BA0003C16

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:
The RP23-3C16 BAC Library has been constructed by Kazutoyo Oseagawa and Minako Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone.

FEATURES
source
Location/Qualifiers
1..173131

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="1"
/map="1"

```
/clone="RP23-3C16"
repeat_region /rptc_family="Alu" 1084.1097
repeat_region /rptc_family="B4" 1098.1194
repeat_region /rptc_family="ID" 1643.1708
repeat_region /rptc_family="B2" 2034.2233
repeat_region /rptc_family="B2" 2401.2762
repeat_region /rptc_family="MALR" 3195.3410
repeat_region /rptc_family="L2" 3450.3544
repeat_region /rptc_family="MALR" 4140.4860
repeat_region /rptc_family="L1" 4837.4891
repeat_region /rptc_family="ERV1" 5137.5192
repeat_region /rptc_family="L1" 5594.5658
repeat_region /rptc_family="ERV1" 6208.6606
repeat_region /rptc_family="ERVK" 6754.7097
repeat_region /rptc_family="ERVK" 7191.7410
repeat_region /rptc_family="L1" 7415.7513
repeat_region /rptc_family="L1" 9015.9225
repeat_region /rptc_family="B2" 9683.9836
repeat_region /rptc_family="B4" 9708.9843
repeat_region /rptc_family="Alu" 9876.9990
repeat_region /rptc_family="B2" 10126.10319
repeat_region /rptc_family="L2" 10456.10637
repeat_region /rptc_family="B2" 12237.12655
repeat_region /rptc_family="L1" 12648.13442
repeat_region /rptc_family="L1" 13445.13614
repeat_region /rptc_family="ERV" 13452.13585
repeat_region /rptc_family="ERVK" 14087.14294
repeat_region /rptc_family="B4" 14257.14624
repeat_region /rptc_family="MALR" 14622.14689
repeat_region /rptc_family="B4" 17030.17193
repeat_region /rptc_family="B4" 17231.17277
repeat_region /rptc_family="B4" 18109.18231
repeat_region /rptc_family="B4" 18757.18982
repeat_region /rptc_family="MALR" 19152.19291
repeat_region /rptc_family="Alu" 19667.19876
repeat_region /rptc_family="B4" 20109.20258
repeat_region /rptc_family="B4" 20984.21106

/rptc_family="Alu" 24485.25438
repeat_region /rptc_family="L1" 25873.25992
repeat_region /rptc_family="L1" 26010.26117
repeat_region /rptc_family="B4" 26042.26154
repeat_region /rptc_family="Alu" 27770.28295
repeat_region /rptc_family="ERV1" 29705.30153
repeat_region /rptc_family="L1" 30464.30599
repeat_region /rptc_family="Alu" 31056.31202
repeat_region /rptc_family="MER1_type" 32828.32936
repeat_region /rptc_family="Alu" 32855.32961
repeat_region /rptc_family="B4" 32988.33382
repeat_region /rptc_family="MALR" 33454.33643
repeat_region /rptc_family="L1" 33651.33765
repeat_region /rptc_family="Alu" 33882.33984
repeat_region /rptc_family="Alu" 34576.34617
repeat_region /rptc_family="B2" 34628.34770
repeat_region /rptc_family="B2" 35256.35305
repeat_region /rptc_family="B2" 35531.35690
repeat_region /rptc_family="L1" 35708.35905
repeat_region /rptc_family="B2" 36174.36385
repeat_region /rptc_family="B2" 38346.38491
repeat_region /rptc_family="Alu" 38961.39070

Query Match 10.8%; Score 541.8; DB 10; Length 173131;
Best Local Similarity 68.3%; Pred. No. 1.5e-92;
Matches 977; Conservative 0; Mismatches 372; Indels 82; Gaps 13;

QY 1787 ATGTCAATCCCTAATAACTATATATACCAATATGACACTTAATTGTGGCCAGAAACAGA--- 1843
DB 115748 ATGCACATTCGTGGAAAGCAATGATATCCACAGAACTGTGTGGAGTGAAGAG 115807
QY 1844 ----ACCTTAACGTGGCCAAATTTATTTATTCATTAATAACAGCTGCTGTGTTTCACTTG 1899
DB 115808 ACCTTACTCAGATGACCAGATTATTT-CTCGGACAGACCTCTGTTTGTAGTGTG 115866
QY 1900 TGCACATCTGAATGAGAGCAATCCCTGTCTGATGTGAGTTTCTTGCACTGATTAAGGANA 1959
DB 115867 TGCACACATGAGCGGTAGC--TGTGCGGTTCTGATGATTTCTTGCACTGATGAGCTTA 115924
QY 1960 AACTGCTGAAGTTGTGAGGCTCTCCAGGACAGAGCCATCATGTGATCATTAAGAAAGCTTC 2019
DB 115925 AACTGCTGAAGTCATGAGACTCTCCAGGACAGAGCCCTGTGTGATGTACAGAGAAAGCTTC 115984
QY 2020 CAGCGCTGACCTCTGCGCAAAAGAGAGAAAGAGTAAGAGGAGGAGGAGGAGAA 2079
DB 115985 CTCTCTGCTGACCTCTGCGAGAGGAGGAGACAAA-----TGAGAAAGAGCGGAGTGGAGAA 116039
QY 2080 GGTTCAAGTGCAGGATTTTCTCTTGAACCTACAGATTATGGGTCAAGAGCTGTGTGCAA 2139
DB 116040 GGTTCAAGGCGGATTTTCTCTTGAACCTGAGTATGGGTCAAGAGCTGTGTGCAA 116099
```


QY 2140 AGAGCTGACGCTGGATGCACTGCTACCATTTGTCAGAGGAGGCGAGCAACAGCT 2199
 DB 116100 AGAGGCTACAGCCAGAGATGCAAGTTGCTACCAACGTTGAGAGGAGGAGCGCAAGCT 116159
 QY 2200 GTGCGAGGAGTCAAGCTGAGAGACCAAGGAG---GCTGCGTTCAAGATTTGATTTCCACAG 2256
 DB 116160 GTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 116219
 QY 2257 AGAGAGCATCTTCTCTATGATTAACCAAGAGTTCAAGACTC---CCTTTTGTAC 2313
 DB 116220 AG-GAGCATCTTCTTATATAGTAACCAAGCAATTCAGACAGACACTCCCTTTTGTAC 116278
 QY 2314 GGGATCTGATTTCTTCCGCTGAGTCTAAACCAATTAATTAATTAATTAATTAATTAATTAAT 2373
 DB 116279 AGAGCTCTTATCTCTGCGGAGTCTAACTCATTAAGAAATTTCTATTAAGTCAAC 116338
 QY 2374 AGAAATTTATGCTGATGATTTATCAATTTGGGGAATTTCTGTTAAACCAAGGAGGAG 2433
 DB 116339 AGAGATTTATGATTTATGATTTATCAAGTTGGGAGATGATCTTTAAACCAAGGAGGAG 116398
 QY 2434 ATATCTCTGCTGCTTGGGCTGCAAGAACTCACTTGGCTTGAAGTCAAGAACTGAG---T 2490
 DB 116399 ATATGCTCTGCTTGGGCTGCAAGAACTCACTTGGCTTGAAGTCAAGAACTGAGTTCT 116448
 QY 2491 TCTCTCAAAATCTCTAAGTCTCTAATTTAAGAGCTGAAACTTAAGAGGAGCACTGCAAGT 2550
 DB 116449 TCTCTGACCTCTGAGCCATTAATTTGAGGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAG 116508
 QY 2551 ATATGCTGATGCTGATGATTTGATTTGAACTTAATTAATTAATTAATTAATTAATTAATTAAT 2610
 DB 116509 ACCGATAGTCCGCTGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 116568
 QY 2611 CATGATTTATTTCCCTTACAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2670
 DB 116569 CTATCTAATTAATTTATTTAGGAGCTTATTAAGTCAAGAGAGGAGGAGGAGGAGGAGGAG 116628
 QY 2671 AACAGAAAAGTTGCAAAATCATCAAGAGATTAGTTAACTGAAATGTTAAAG 2730
 DB 116629 AAATCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 116678
 QY 2731 CACATTTTATGATGAG 2790
 DB 116679 AAATCTGATGATGAG 116738
 QY 2791 TCAGGAT--CTTCAAGAAAAGTCTTTGGGAGGAGTCAAGAGAGAGAGAGAGAGAGAGAG 2848
 DB 116739 AGAGATCTGATTTCAAGAAAAGTCTTTCAAGGAGGAGTCAAGAGAGAGAGAGAGAGAG 116798
 QY 2849 AAGAGTAAATGAGCAAAATTAAGAAA-----AAATGAT 2883
 DB 116799 ACCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 116858
 QY 2884 TTATACGAGCACTGATCTTTTACCTTTGATTAATTTATTTTCTCTTTCTTTTCTTTCTTTCTC 2943
 DB 116859 TTTCGCGCATCATCATTTGATCTTTGATCAATTTTGTCTGCTGATGATTTCTTT----- 116912
 QY 2944 TAGCTTAACAGACCTTAAG 3003
 DB 116913 ----TCTCTAGCTTAACAG 116968
 QY 3004 CAAGCTGCGCTGCGCAAGAGTCAAGAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3063
 DB 116969 GAGGCTGCGCTGCGCAAGAGTCAAGAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 117028
 QY 3064 CAGAGAGAGAGTCTGAGGAG 3123
 DB 117029 CAGAGAGAGAGTCTGAGGAG 117088
 QY 3124 AAGAGATTTATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3174
 DB 117089 GAGAGATTTATGATGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 117139

RESULT 5
 AC128908
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-320D4, WORKING DRAFT SEQUENCE, 2
 unorderded pieces.
 AC128908
 AC128908.3 GI:25073564
 HTG, HTGS_PHASE1, HTGS_DRAFT; HTGS_FULFROP.
 Rattus norvegicus (Norway rat)
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 179341)
 Muzny,D,Marie, Metzker,M,lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Albrooks,S, Amin,A, Angiano,D, Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biewald,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Cesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denison,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durkin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlik,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M, Hollins,B, Howells,S, Hulik,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A, Karpach,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,Y, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lounsbury,L, Louissege,H, Lozano,R,J, Lu,X, Ma,J, Maheshwari,M, Mahindaratne,M, Mahmood,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenen,B, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwakoelameh,O, Okwomu,G, Olarunmugun,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkuch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Fu,L, Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Register,S, Riggs,F, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H, Shetty,J, Shvartbeyn,A, Sison,I, Sitter,C,D, Smaj,D, Sneed,A, Sodergren,E, Song,X,Z, Sorelle,R, Sosa,J, Stemle,M, Strong,R, Sutton,S, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tiney,A, Trejos,Z, Uman,K, Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wooden,H, Worley,K, Williams,G, Willson,R, Wleczyk,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,D, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstein,G, and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 179341)
 Worley,K.C.
 Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 179341)
 Rat Genome Sequencing Consortium.

REFERENCE
 AUTHORS

Db 21716 GTCCGCTTCTTCTCTCCGCTTAACAGCTTAAAGAGCATGCTTCCTGCT 21775
Qy 2983 TCATTCACCCCGAGGGAATTCAGAGTGCCTCCGACCAAGATTCAAGATAAAGGCA 3042
Db 21776 TCATTCACCCCGAGGGAATTCAGAGTGCCTCCGACCAAGATTCAAGATAAAGGCA 21835
Qy 3043 CGAGTAACAGTATGTTTTCAGAAAGAGATCTGAGGAGGCTTTTCAGTGGTGAAT 3102
Db 21836 CGAGTAACAGTATGTTTTCAGAAAGAGATCTGAGGAGGCTTCATGGTGAAT 21895
Qy 3103 GACACGTGATGTTGTAACAGAGATTAAATGAGCTTTGTCTAGGAGGCCGCCCTT 3162
Db 21896 GAACCGCTGACAGGACATCAAGAGATTATGAGTGTCTAGGAGAGCTAGTCTAG 21955
Qy 3163 GCTTC 3167
Db 21956 CCTTC 21960

RESULT 6
LOCUS G50119 431 bp DNA linear STS 30-MAR-2000
DEFINITION SHGC-83058 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G50119
VERSION G50119.1 GI:5221446
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 431)
AUTHORS Olivier, M. and Cox, D.R.
TITLE Unpublished, Olivier, M., Cox, D.R. (2000)
JOURNAL Unpublished (2000)
COMMENT

Contact: Michael Olivier, David R. Cox
Stanford Human Genome Center
Stanford University School of Medicine
4005 Miranda Ave. 2nd Fl., Palo Alto, CA 94025, USA
Tel: (650) 320-5800
Fax: (650) 320-5801
Email: olivier@shgc.stanford.edu
Primer A: CCTCCCTTAATCAATGACG
Primer B: TCCAACTTCCCACTTGACTTA
STS size: 290
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes
Denaturation: 94 degrees C for 30 seconds
Annealing: 60 degrees C for 30 seconds
Polymerization: 72 degrees C for 23 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9700
Protocol:

Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
AmpliTaq Gold Polymerase: 0.07 units/ul
Total Vol: 5 ul
Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 10 mM
pH: 8.3

FEATURES
Source
BAC ends sequenced at TIGR from the RPC11 BAC library. Designed and developed at the Stanford Human Genome Center.

Location/Qualifiers
1..431
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

/map="2"
/clone_lib="Human"
STS
primer_bind 137..426
primer_bind complement (404..426)
ORIGIN

Query Match 8.0%; Score 401.4; DB 11; Length 431;
Best Local Similarity 98.5%; Pred. No. 8.4e-66;
Matches 405; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3647 ATTGAGAAATTCAGAAAGAAATAGACATTCCTACTGACCCAGAGCAATATAC 3706
Db 21 ATCCGCGAATTCAGAAAGAAATAGACATTCCTACTGACCCAGAGCAATATAC 80
Qy 3707 TCTTCAGGCTGTGGAGGACACAGTGGTAAGCGTCTTAACAGGTTTTTAATCCCT 3766
Db 81 TCTTCAGGCTGTGGAGGACACAGTGGTAAGCGTCTTAACAGGTTTTTAATCCCT 140
Qy 3767 CCTTAATCAATACAGAGTTTGTATGCAACCTGGAATTCCTCATTCCTC 3826
Db 141 CCTTAATCAATACAGAGTTTGTATGCAACCTGGAATTCCTCATTCCTC 200
Qy 3827 CACTGCGCTTATAGAGAAAGTGAAGTGGTTCGCAATTATGTATGCAAG 3886
Db 201 CACTGCGCTTATAGAGAAAGTGAAGTGGTTCGCAATTATGTATGCAAG 260
Qy 3887 ATGATTAATCTAGATTTTATATTTGCAAAATACAAATATCTGAGATTAAT 3946
Db 261 ATGATTAATCTAGATTTTATATTTGCAAAATACAAATATCTGAGATTAAT 320
Qy 3947 ACTGCTTATCCAAAGCTAAGTAACTTTGTTGTAACCAACTTTGTTAATATAG 4006
Db 321 ACTGCTTATCCAAAGCTAAGTAACTTTGTTGTAACCAACTTTGTTAATATAG 380
Qy 4007 TAAAGATTCAGTAATCCCTTTTATGCAAGTGGGAAAGTGGATGTC 4057
Db 381 TAAAGATTCAGTAATCCCTTTTATGCAAGTGGGAAAGTGGATGTC 431

RESULT 7
LOCUS AL451085 182166 bp DNA linear PRI 09-MAR-2002
DEFINITION Human DNA sequence from clone Rpl1-307C12 on chromosome 1, complete
ACCESSION AL451085 AC027440
VERSION AL451085.20 GI:19309454
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 182166)
AUTHORS Brown, A.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Mar 8, 2002 this sequence version replaced gi:19031396.
During sequence assembly data is compared from overlapping clones.
When differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given

Consensus quality: 191604 bases at least Q30
Consensus quality: 192567 bases at least Q20
Insert size: 178000; agarose-fp
Insert size: 194189; sum-of-contigs
Quality coverage: 6.28 in Q20 bases; agarose-fp
Quality coverage: 5.78 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 11 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2708: contig of 2708 bp in length
2709 2808: gap of unknown length
2809 6955: contig of 4147 bp in length
6956 7055: gap of unknown length
7056 14453: contig of 7398 bp in length
14454 14553: gap of unknown length
14554 23402: contig of 8849 bp in length
23403 23502: gap of unknown length
23503 37767: contig of 14265 bp in length
37768 37867: gap of unknown length
37868 54470: contig of 16603 bp in length
54471 54570: gap of unknown length
54571 82102: contig of 27533 bp in length
82103 82202: gap of unknown length
82203 118832: contig of 36630 bp in length
118833 118932: gap of unknown length
118933 191118: contig of 72186 bp in length
191119 191218: gap of unknown length
191219 192795: contig of 1577 bp in length
192796 192895: gap of unknown length
192896 195189: contig of 2294 bp in length.

FEATURES

source

1..195189
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-163D16"
1..2708
/note="assembly_name:Contig10"
2809..6955
/note="assembly_name:Contig11"
7056..14453
/note="assembly_name:Contig12"
14554..23402
/note="assembly_name:Contig13"
23503..37767
/note="assembly_name:Contig14"
37868..54470
/note="assembly_name:Contig15"
vector_end:right
54571..82102
/note="assembly_name:Contig16"
82203..118832
/note="assembly_name:Contig17"
118933..191118
/note="assembly_name:Contig18"
191219..192795
/note="assembly_name:Contig19"
192896..195189
/note="assembly_name:Contig19"
clone_end:SP6
vector_side:left"

ORIGIN

Query Match 7.7%; Score 383; DB 2; Length 195189;
Best Local Similarity 66.6%; Pred. No. 1.7e-62;

Matches 676; Conservative 0; Mismatches 315; Indels 24; Gaps 8;

QY 5 GGCACATCTGTGCTCACGGCTGATTTCCACAGCACTTTGGAGGC-----AGAGAT 58
Db 188924 GGCACACCTGTGTGCTCACACCTGTAATCCAGATTTTGGAGGCTTAAGAGGACAT 188983
QY 59 CGCTTAGCTCAGGAATTGAGACAGAGCTACGTAACTAGTGAACCTGTCTGTACA 118
Db 188984 CACTTAGAGTCAGAGATTCAGAGCCAGCTGCGCAACATGTTGAATCCCGTCTATTT 189043
QY 119 AATAATTAAGAATT-TTCAGGATGTGCGTGCACCCCACTGCTATTTGGGA 177
Db 189044 AAAAAATCAAAAATTAGCTGGTGCAGTGCCTATGCTTAATCCAGACATTTGGGA 189103
QY 178 GGCTAGAGTAGAGGAATGCTTGAAGCCAGAGTTGAAGAGAAGCTTAGCAACATAGT 237
Db 189104 GGCTAGAGTAGGTGATCACCTGAGGTAGAAGTTGAACAGAGCTTGGCAACATGTTG 189163
QY 238 AGACCTGTGTCTA-TAAAAAATAATATAGCTGTGTCTTGGCACAGGCTTGA--GCTA 294
Db 189164 AAACCCATCTTACTTAAACAAATTTAGCCGGGTGTGTGGCACAGCTGTATCCCA 189223
QY 295 GCTACTCGAAGACTGAGGTGAGGATCAC-TGAGCCAGAGGCTGAGCTGACATGA 353
Db 189224 GCTACTCGGAGGCTGAGGAGGAGATCACTTGAACCTGGAGGAGGAGTTGACATGA 189283
QY 354 AAGAGATACCCAGCTGATTCAGCTTGAAGAC-AGAGGAGACCTGTTCCAAA 412
Db 189284 GCTGAGATCACCTTGTGACTTCGTTGGCAAGAGTGAACCTGTCTAATAA 189343
QY 413 AAAAAAAAAAAAAAAAAATGCAAGAAAAACATCATAACTTGACCTGGACATTACTTT 472
Db 189344 AATTAATAATTAATTTAGCCGGGTGTGTGTGCTGTATCCCTGCTACAGAGGCTG 189403
QY 473 ATGTAGTAATTCACAACTTTTGAAGAAATTTAGCATTTCTATTAATATTTATTA 532
Db 189404 AGGTGAGAGATCGCTTGAACCCGAGAGGACAGAGGTGAAGGACCAAGTACACAC 189463
QY 533 ATTATTTATTTAATTAATTAATGAATTAATTTTGAAGAACTAGCTT----- 583
Db 189464 TGCACTCAGCTGGGCGAGAGAGCTGTCTCAATTTAAAAAACAATTCACAGAA 189523
QY 584 --CTCACTCTCAGTGTCACTCAAACTTTAATGTCTTTGGCCGGGTGCGTGCCTC 641
Db 189524 TAAATTACTTCCCTCTAGCTAAACAAACAGTAAGTAAAGTGGCGGGCGTGCCTC 189583
QY 642 ACGCTGTAATCCAGCACTTTGGAGGCGGAGCGGCTGATCAAGTTAGAGATC 701
Db 189584 ACGCTGTAATCAAGCACTTTGGAGGCGGAGCGGCTGAGGATCAAGTGAAGATC 189643
QY 702 GAGACCATCTGGCTTAACAGGGAACCTGCTCTACTAATAATAC-AAAAAATTAGC 760
Db 189644 GAAACCATCTGCTTAACAGGGAACCTGCTCTACTAATAATAAAAAATTAGC 189703
QY 761 CGGATCGGTGTCAGACGCTGTAGTCCAGCTGCTCAGAGGCTGAGGACAGAAATGC 820
Db 189704 TGGCAAGGCGGTGCGCTGTAAGTCCAGCTAATCCGAGGCTTAAGGACAGAAATGC 189763
QY 821 TGTGAACCGGAGGCGGAGCTTTCAGTGAAGCCGAGATTTGCCCACTGACCCAGCCTC 880
Db 189764 CGTGAATCGGAGGCGGAGCTTTCAGTGAAGCCGAGATTTGCCCACTGACCCAGCCTC 189823
QY 881 GCGACAGGCGGAGCTGCTCAAAAAAAGTTGAATGATGCTTTGA 940
Db 189824 GCGACAGGCGGAGCTGCTCAAAAAAAGTTGAATGATGCTTTGA 189883
QY 941 GCGAATGATCTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 995
Db 189884 GGAATTAATCTCCCATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 189938

RESULT 9
AC104531/c

LOCUS	AC104531	16195 bp	DNA	linear	PRI 14-JUL-2002
DEFINITION	Homo sapiens chromosome 19 clone LTNLR-266A7, complete sequence.				
ACCESSION	AC104531				
VERSION	AC104531.3	GI:21747474			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 16195) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission Unpublished				
AUTHORS	2 (bases 1 to 16195) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission				
JOURNAL	Submitted (13-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
REFERENCE	3 (bases 1 to 16195) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission				
AUTHORS	Submitted (20-MAR-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
JOURNAL	4 (bases 1 to 16195) DOE Joint Genome Institute and Stanford Human Genome Center. Direct Submission				
REFERENCE	Submitted (14-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
AUTHORS	On Jul 14, 2002 this sequence version replaced gi:19551122. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov				
JOURNAL	Finishing Completed at Stanford Human Genome Center www.sngc.stanford.edu				
COMMENT	Quality: Phrap Quality >=40 99.1% of Sequence; Estimated Total Number of Errors is 0.1. NOTE: This insert is not the entire sequence of the clone. It is clipped at the overlaps with AC010331 and AC008985. The number of bases overlapped with AC010331 is 9402 and with AC008985 is 6514.				
FEATURES	Location/Qualifiers				
source	1..16195 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="19" /clone="LTNLR-266A7"				
ORIGIN	4				
Query Match	7.6%; Score 380.6; DB 9; Length 16195;				
Beet Local Similarity	70.5%; Pred. No. 5.8e-62;				
Matches 590; Conservative	0; Mismatches 219; Indels 28; Gaps 5;				
OY	103	AACCTGTCTGTCAATAATAAAGATTTTCCAGCAGTGGGGCGTCA	CCCCAGT	162	
Db	5648	AACCTTGTGTTGAAAAAATCTTGAATATGAAACAGGTGACAGTGCATGCTGCAT	55898		
OY	163	GCCAGCTATTTGGGAGGCTGAGTAGAGAGATGCTTAAGCCAGATTTAAGCAAGC	222		
Db	5588	CCGACACTTTTGGGAGGCCCAAGGACGTGGATACCTTAGGCCCAAGATTTCAAGCAAGC	5529		
OY	223	CTAGGCAACATAGTGAGACCTGTGTCT-----ATAAAAATATTAAGCTGGTTG	272		
Db	5528	CTGGGCAACATAGTGAGACCTGTGTCTCTTACAAAAAGTAAAAAAAATTTAGCCACATG	54659		
OY	273	TCTTGGGCAAGGCCGTCGAG--CTAGCTACTGGGAAGTGAAGTGGGGAGATCAAC-TCAG	329		
Db	5468	TGTGTGACGCTCTCTGTATGTCCTCCGCTTACTGTATGAGTGAAGTGGGAGATCACTTAG	54098		
OY	330	CCGAGAGGCTGAGGCTGACAGTGAACAGTGAATCCACAGCTGSAATTCAGCTTGAAAC	389		
Db	5408	CCGAGAGGCAAGAGGTTCAGTGAAGCAAGATTCAGCACTGCACTCCAGCTGGGCGAC	53498		
OY	390	AGAGGAGACCTGTGTTCCAAAAAATTTAAAAAATGCAAGAAAGATCATTA	449		

Db	5348	AGACCAGACCCCGGCTCCAAAAA	AAAAAAAAAAAAAATCCTGAAATGACAGATGATG	5289
Qy	450	ACTT--GACCTGGAGCA	TAACCTTTATGTATGATAATTCACATCTTTAGAGAAAT	507
Db	5288	TTTGACAAACATTATGAAAGTAC	TTAGTCCACCAACTGTATCTTTAAAAATGGTTCAT	5228
Qy	508	AGCATTTCTGATATAATGTAT	TTTAAATTTATTTATTTAAATTCAAATGCAATTTAAAT	567
Db	5228	TTTATGTATGATATTTGTA	CCACATATAATTTAATTAACAGCAGCTAACATTTATGTT	5169
Qy	568	TCGTGAAAACTAGCTTCT	CACTCTCTCAATTTGTACGCAAACTTTAATGTCCTTTGGCC	627
Db	5168	GC-----	TTTTGATGATCAGGCATTATTTCAAAAGCTTTACGTATGTGCC	5122
Qy	628	GGGTCGGTGAGCTCACG	CTGTAAATCCACACACTTTTGGAGCGAGCGGGTGATATAC	687
Db	5121	AGAGAGCGGTGGCTACG	CGCTGTATCCACAGACTTTTGGAGGCCAAAGCGGGCGGATATAC	5062
Qy	688	AAGGTTAGAGATGTGA	AGACATCTTGCTTAAACACGCTGAAACCTGTCCTTACTTAAAT	747
Db	5061	GAGGTTCAGGAGATTCG	AACCATCTTGGCTTAACACAGTAAACCCCATCTCTACTTAAAT	5002
Qy	748	ACAAAAAATTAGCCGG	TGGCGGTCCAGAGCGCTGTAGTCCACAGCTGCTCAGAGGCTGA	807
Db	5001	ACAAAAAATTATACCG	AGCATGTGTGGCGGTGCTCTGCAGTCCCACTTACAGAGGCTGA	4942
Qy	808	GGCAGAGAAATGCTG	TAACCCGGGAGCGGAGCTTGCAGTGAGCCGAGATTTGGCCACT	867
Db	4941	GGGAGGGAATTGGGT	TAAACCCGGGAGGACAGAGCTTTCAGTGAAGCCGAGATTCGGCCACT	4882
Qy	868	GCACTCCAGCTGGG	GCAGATGGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA	924
Db	4881	GCACTCCAGCTGGGT	ACAGAGGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA	4825

LOCUS	AC008753	147330 bp	DNA	linear	PRI 22-MAR-2003
DEFINITION	Homo sapiens chromosome 19 clone CTD-302206, complete sequence.				
ACCESSION	AC008753				
VERSION	AC008753.9	GI:25140998			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 147330)				
TITLE	DOE Joint Genome Institute and Stanford Human Genome Center.				
JOURNAL	Direct Submission				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 147330)				
TITLE	DOE Joint Genome Institute.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
AUTHORS	3 (bases 1 to 147330)				
TITLE	DOE Joint Genome Institute and Stanford Human Genome Center.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (29-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
AUTHORS	4 (bases 1 to 147330)				
TITLE	DOE Joint Genome Institute and Stanford Human Genome Center.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (06-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
AUTHORS	5 (bases 1 to 147330)				
TITLE	DOE Joint Genome Institute and Stanford Human Genome Center.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (21-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
AUTHORS	6 (bases 1 to 147330)				
TITLE	DOE Joint Genome Institute and Stanford Human Genome Center.				
JOURNAL	Direct Submission				

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Thomas, S., Uemami, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleciyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 175826)
Worley, K.C.
Submitted (22-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 9, 2000 this sequence version replaced gi:9719697.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: HMZB

Center clone name: RP11-498A2

Sequencing vector: M13: L08821

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 144734 bases at least Q40

Consensus quality: 162873 bases at least Q30

Consensus quality: 169424 bases at least Q20

Estimated insert size: 170770: sum-of-coverage estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 3.3x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 22 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 21031: contig of 21031 bp in length
21032 21131: gap of unknown length
21132 40475: contig of 19344 bp in length
40476 40575: gap of unknown length
40576 57617: contig of 17042 bp in length
57618 57717: gap of unknown length
57718 75708: contig of 17991 bp in length
75709 75808: gap of unknown length
75809 89018: contig of 13210 bp in length
89019 89118: gap of unknown length
89119 100348: contig of 11230 bp in length
100349 100448: gap of unknown length
100449 110704: contig of 10256 bp in length
110705 110804: gap of unknown length
110805 123156: contig of 12352 bp in length
123157 123256: gap of unknown length
123257 132467: contig of 9211 bp in length
132468 132567: gap of unknown length
132568 141461: contig of 8894 bp in length
141462 141561: gap of unknown length
141562 146576: contig of 5015 bp in length
146577 146676: gap of unknown length
146677 151301: contig of 4525 bp in length
151302 151301: gap of unknown length
151302 155722: contig of 4421 bp in length
155723 155823: gap of unknown length
155823 160148: contig of 4326 bp in length
160149 160248: gap of unknown length
160249 163024: contig of 2776 bp in length
163025 163124: gap of unknown length

FEATURES

source

1. 175826
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-498A2"

ORIGIN

Query Match 7.6%; Score 378.2; DB 2; Length 175826;

Best Local Similarity 67.0%; Pred. No. 1.4e-61;

Matches 665; Conservative 0; Mismatches 303; Indels 25; Gaps 8;

5 GGCCATGCTGTGCTCAGCGCTGATTCGCCAGCACTTGGAGGC-----AGAGGAT 58
161959 GGCACACGCGGCTGCACACCTGATCCAGCACTTGGAGGCTAAGAGGCGAGAT 162018
59 CGCTGAGCTCAGGAATTGGAGACAAAGCTAAGTAACTAGGAACCTCTGCTGATA 118
162019 CACTTGAAGTCAGGAGTTGAGACCAAGCTGCGCAACATGGAAATCTCCCTCTATT 162078
119 AATATTAAGAAATTT--CCAGCATGCTGCGCTGCAACCCAGTGCAGCTATTGGG 176
162079 AAAAAATCAAAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 162138
177 AGCTGAGGTAGAGGAATGCTTGAAGCCAGAGATTGAAGACAGCTTGGCAACATGT 236
162139 AGCTGAGGTAGAGGTGATGATCAGCTGAGTCAGAAAGTTGAACCAAGCTGCGCAATGCT 162198
237 GAGACCTGTGCTA-TAAAAAATTAATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 293
162199 GAAACCCATCTCTACTAATAAATAATTAAGCGGCGTGGGCAACGCTGTATATCC 162258
294 AGCTACTCGAAGACTGAGGTGGAGATCAC-TGAGCCAGAGAGGCTGAGCTGAGTG 352
162259 AGCTACTCGAAGAGGCTGAGGCAAGAAATCACTTAACCTGGAAGGCAAGGTTGAGTG 162318
353 AACAGTATCACCCAGCTGATTCAGGCTGGAAGAC-AGAGGAGACCTGTTCCAAA 411
162319 AGCTAGATCACGCTATGCACTTCAAGCTGGGCAACAAAGATGAATCTCTCTCAAAA 162378
412 AAAAAAAAAAAAAAAAAATGCAAGAAAGACATCTAACTGACCTGGAGATATACCTT 471
162379 AAAAAAAAAATAATTAATTAAGCGGCGGTGTGTGCTGTATCTCTGATCAGAGAGCT 162438
472 TATGTATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 531
162439 GAGGTAGAGATTCCTTGAACCCGAGAGGAGAGAGTGAAGTGAAGTGAAGTGAAGT 162498
532 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 581
162499 CTGCACTAGCGCTGGGCAAGAAAGCTGTCTCAATTAATAAAAAAAAAAATTCACAGA 162558
582 -TTTCATCTCTCTGAGTGTGATCAAAATTAATTAATTAATTAATTAATTAATTAAT 640
162559 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 162618
641 CAGGCTGTAATCCAGCACTTGGAGGCGAGGCGGAGTGAATCAAGATTAGAGAT 700
162619 CAGGCTGTAATCAAGCACTTGGAGGCGAGGCGGAGTGAATCAAGATTAGAGAT 162678


```

repeat_region      /rpt_family="TTTTTC)n"
                    /evidence=not experimental
                    /rpt_family="TTCCTC)n"
repeat_region      complement(18631..18867)
                    /evidence=not experimental
                    /rpt_family="AluB0"
repeat_region      complement(19046..19344)
                    /evidence=not experimental
                    /rpt_family="AluBp"
                    19436..19456
repeat_region      /evidence=not experimental
                    /rpt_family="AluB_rich"
                    complement(19457..19576)
repeat_region      /evidence=not experimental
                    /rpt_family="FLAM_C"
                    20700..20991
repeat_region      /evidence=not experimental
                    /rpt_family="AluSx"
                    21256..21373
repeat_region      /evidence=not experimental
                    /rpt_family="FLAM_C"
                    21729..21846
repeat_region      /evidence=not experimental
                    /rpt_family="L2"
                    22372..22496
repeat_region      /evidence=not experimental
                    /rpt_family="LTR41"
                    22899..22973
repeat_region      /evidence=not experimental
                    /rpt_family="LTR50"
                    23184..23207
repeat_region      /evidence=not experimental
                    /rpt_family="T(G)n"
                    23412..23534
repeat_region      /evidence=not experimental
                    /rpt_family="L2"
                    complement(23616..23708)
repeat_region      /evidence=not experimental
                    /rpt_family="MIR"
                    27023..27322
repeat_region      /evidence=not experimental
                    /rpt_family="AluBb"
                    27460..27765
repeat_region      /evidence=not experimental
                    /rpt_family="AluSx"
                    28489..28748
repeat_region      /evidence=not experimental
                    /rpt_family="AluSx"
                    28768..28796
repeat_region

Query Match      7.5%; Score 376; DB 9; Length 162728;
Best Local Similarity 68.8%; Pred. No. 3,7e-61;
Matches 640; Conservative 0; Mismatches 265; Indels 25; Gaps 8;

QY 12 GCTGTGCTCAACGCTGTATTCACGACTTGTGGAGGCGG-----GAGGATCGCTGA 65
DB 49333 GCAGTGGCTTAAAGCTGTATTCGACACTTGGAGGCTGAGCCGACGACGATGCTTGA 49274
QY 66 GCTCAGGAATTTGAGACAGACCTTACATAGTGAACCTGTGTACAAATATA 125
DB 49273 GCTCAGAGTTTCAAGACAGCAGCTGTCAATATGTGAGACC-CTGTCTTCAAAAAATA 49215
QY 126 AAAAATTTTCCAGGATGTGGCGTGACACCCCAAGTCCAGCTATTGGAGGCTGAGG 185
DB 49214 CATGAATTTTCCAGGCGCGGTGCTACGCTTGAATCCAGACATTTGGAGGCGCAAG 49155
QY 186 TAGGAGGAATGCTTGAAGCAGAGTTGAAGACAGCCTAGGCAATAGTAGACCTG 245
DB 49154 CAGGCAATACCTGAGGTGGAGTTCAAGACAGCCTACCAACATGAGAAATCCA 49095
QY 246 TGTCTATTAATAAAT---AATTAGCTGTGTGCTTGGACAGGCTTGA--GCTAGCTAC 299

```

```

DB 49094 ACTCACTTAATAAATAAATAATTTAGCCGGGATGTGTGATGCTGTATTCACGACTAC 49035
QY 300 TCGAAGAAGCTAGGTGGAGGATGAC-TGAGCCCAAGAGGCTGAGGCTGACGTGAACGT 358
DB 49034 TCGGAGAGCTAGGAGGAGAGAAATTAATTGAACCTGGAGGCGGAGGTTGACGTGCGCGA 48975
QY 359 GATCACCCAGCTGATTTCCACCTCGGAMAGACAGAGGGGAGACCCGTTTCAAAAAAANA 418
DB 48974 GATTGCGGCATTGCACTCCACCTGGGTGAC-AAGGCAAACTGTGTCTCAAAAAAANA 48916
QY 419 AAAAAAAAAAATGCAGAAAGAAACATCAATACTTGAACCTGGGACATTAATTTATGTA 478
DB 48915 AAAAAAAAAAATGCAAT-ATTAGCCAAAGGTGTGGCAGCATGCTGTACTGCA 48862
QY 479 TGAATTTCACAATCTTTTGAAGAGAAATTTAGCATTTTGTGAATTAATTAATTTA 538
DB 48861 GCTACTCGGAGGCTGAGGTGGAGGTGTGGCTAGAGCCAGGAAGTTGACGTAGCTAGA 48802
QY 539 TTATTAATAATTTCAAAATGAATTAATTTCTGAGAACTAGCTTCTCACTCTCAG-- 596
DB 48801 TCATGCCACTGCACTCCAGCCTGGGCGACAGTCAGACATGCTTCAAAAGAAAAAAG 48742
QY 597 --TTGTCAATCAAACTTTAAATGTCTTTGGCCGGGTGCGGTGCTCAGCCTGTATTC 654
DB 48741 AAGTTGCTTAAAGAGAGGCAAAATATCTGGCGAGGCGCAGGTGCTCAGCCTGTATTC 48682
QY 655 CAGCACTTTGGAGGCGGAGGCGGCTGATCAGAGGTTAGAGATTCAGAGCCATCTCGG 714
DB 48681 CAGCACTTTGGAGGCTGAGGCGGCTGATCAGAGGTTAGAGATTCAGAGATTCAGG 48622
QY 715 CTAACACGGTGAAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 774
DB 48621 CTAACACGGTGAAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 48562
QY 775 GAGCGCTGTATGCTCCAGCTGCTCAGAGGCTGAGGCGAGAGATGTGTAACTCCGGAG 834
DB 48561 GGCACCTGTATGCTCCAGCTGCTCAGAGGCTGAGGCGAGAGATGTGTAACTCCGGAG 48502
QY 835 GCGGAGCTTGCAGTACGCGAGATTTGCCCACTGCACTCCAGGCGGCGAGAGTGGAG 894
DB 48501 GTGAGCTTGCAGTACGCGAGATTTGCCCACTGCACTCCAGGCGGCGAGAGTGGAG 48442
QY 895 ACTGTGTCAAAAAAAAAAAAAAAAAAAAA 924
DB 48441 ACTCACTTAATAAATAAATAATTTAGCCGGGATGTGTGATGCTGTATTCACGACTAC 48412

RESULT 14
AC022470/c 111469 bp DNA linear HTG 25-JUL-2001
LOCUS Homo sapiens chromosome 6 clone 133K12 map 6p25, WORKING DRAFT
DEFINITION
SEQUENCE, 8 unordered pieces.
AC022470
AC022470.7 GI:15011676
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS McCombie,W.R.
TITLE Human Genomic Sequence, Chromosome 6
JOURNAL Unpublished
REFERENCE
AUTHORS McCombie,W.R.
TITLE Direct Submissiion
JOURNAL Submitted (04-FEB-2000) Lita Annenberg Hazen Genome Sequencing
Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
Harbor, NY 11724, USA
On Jul 25, 2001 this sequence version replaced gi:10440592.
COMMENT
----- Genome Center
Center: Lita Annenberg Hazen Genome Center, Cold Spring Harbor
Laboratory

```

Center code: CSHL
 Web site: <http://www.cshl.org/genseq>
 Contact: mccombie@cshl.org
 Project Information
 Center project name: 133K12
 Center clone name: 133K12

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 33563: contig of 33563 bp in length
 * 33564 34544: gap of unknown length
 * 34545 57280: contig of 22736 bp in length
 * 57281 58260: gap of unknown length
 * 58261 73887: contig of 15627 bp in length
 * 73888 74867: gap of unknown length
 * 74868 85395: contig of 11128 bp in length
 * 85396 86975: gap of unknown length
 * 86976 97534: contig of 10558 bp in length
 * 97534 98513: gap of unknown length
 * 98514 104746: contig of 6233 bp in length
 * 104747 105726: gap of unknown length
 * 105727 109390: contig of 3664 bp in length
 * 109391 110371: gap of unknown length
 * 110371 111469: contig of 1099 bp in length.

FEATURES
 source
 1. 111469
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="6p25"
 /clone="133K12"

ORIGIN
 Query Match 7.5%: Score 374.2; DB 2: Length 111469;
 Best Local Similarity 68.1%; Pred. No. 8.3e-61;
 Matches 637; Conservative 0; Mismatches 263; Indels 35; Gaps 7;

5 GGCACATGCTGTGCTCAGCGCTGATATCCAGCACTTTGG-----AGGAGAGAGAT 58
 |||||
 51735 GGGCGGAGACGTGGCCGATGCTGATCAAGCACTTTGGGCGGTAAAGGGGGTGGAT 51676
 |||||
 59 CGCTTGAGCTCAGGAATTTGAGACAAAGCTTACATATGTAAGAACTCTGTCTGTACA 118
 |||||
 51675 CACCTGAGGCTCTGGAGTTGAGAACCGCTGCGCAAAATGATGTAAC-CACCATCTCTACT 51617
 |||||
 119 AATAATTAAGAAATTTCCAGGATGATGGGCGTGAAGCCCGCCAGTGCACATTTGGGAG 178
 |||||
 51616 AAAAAATCAAAAAGTACCGGTCAGTGGCTCTGCCCTGTATCTTACCACTTTGGGAG 51557
 |||||
 179 GCTGAGTAGAGAGAAATGCTTGAAGCCAGAGTGAAGCAAGCCCTAGGCAATAGTGA 238
 |||||
 51556 GCCGAGGAGCGGAATTTGCTTAAGCATGATTGAACCAAGCTTGGCCAAATGCTGA 51497
 |||||
 239 GACCCTGTCTTATTA-----AATATTTAGCTGTGTCTTGGCACAGGCTTGA-- 290
 |||||
 51496 AACCTTGTCTTATTAACCAAAATAATTTAGCTGGGGGTATGCGCATGCTCTGTAT 51437
 |||||
 291 GCTAGCTACTGGAAGACTGAGGTGGAGATCAC-TGAGCCAGAGAGGCTGAGGCTGCA 349
 |||||
 51436 CCCAGCTACTCGGGTGGCTGGGCGAAGAAATCGCTTGAACCCGAGAGGCGAGGTGA 51377
 |||||
 350 GTGAACAGTATCACCGACGTGATTCAGCTTGAAGACAGAGGAGACCTCTTTTCCA 409
 |||||
 51376 GTGAGCCGAGATCACGCCACTGCGCTTACCTGGGCGACAGAGCAAGACTCATCTCAA 51317
 |||||
 410 AAAAAAAAAAAAAAAAAATGCAGAAAAGACATATAACTTGACTTGGAGCATTAAT 469

Db 51316 AAAAAATATATATATACAAAAT-----TAGCCAGCAAGTGGCAGCAC 51273
 |||||
 Qy 470 TTTATGATGATGAATTCACAACTTTTAGAGAGAATTAAGCATTTCTGTATAAATGAT 529
 |||||
 Db 51272 TGTACTCCAGCTACTTTGGAGCGGAGGACAGAAATCATCTTAACCTGGAGGAGAA 51213
 |||||
 Qy 530 ATAATTATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 589
 |||||
 Db 51212 CTTCAGTAGAGCTGAGCGGACCACTCCTCAGACCGCTGACAGAGCAAGACT 51153
 |||||
 Qy 590 CTTCAGTTGTAGTCAAAACCTTTATGCTTTGGCGGGGGGGTGGCTCAGCGCT 649
 |||||
 Db 51152 CTGTACAT--ACATPACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 51096
 |||||
 Qy 650 AATCCAGCACTTTGGAGGCGCGAGCGGTGATCACAAGTTAGAGATGAGACCAT 709
 |||||
 Db 51095 ATTCACAGCACTTTGGAGGCGCGAGCGGAGATCACAAGTTAGAGATGAGATCAT 51036
 |||||
 Qy 710 CCTGGCTAACACGCTGAAACCTCTCTCTACTTAATAAATTAATTAATTAATTAATTAATTAAT 769
 |||||
 Db 51035 CCTGGCTAACACGCTGAAACCCCATCTCTACTTAATAAATTAATTAATTAATTAATTAATTAAT 50976
 |||||
 Qy 770 TGCCAGAGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGGAGGAGAAATGCTGAACCC 829
 |||||
 Db 50975 TGCCGAGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGGAGGAGAAATGCTGAACCC 50916
 |||||
 Qy 830 GGGAGCGCGAGCTTGCAGTGAAGCCAGATTTGGCCCATCTCAGCTCAGCTGGCGCAACT 889
 |||||
 Db 50915 GGGAGCGCGAGCTTGCAGTGAAGCCAGATTTGGCTGCTCAGCTCAGCTGGCGTGAACAG 50856
 |||||
 Qy 890 GCGAGCTCTGTCTCAAAAAAAAAAAAAAAAAA 924
 |||||
 Db 50855 GCGAGACTCCGTCTCAAAAAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 50821
 |||||

RESULT 15
 AC018410
 LOCUS Homo sapiens chromosome 11, clone RP11-17G12, complete sequence.
 DEFINITION AC018410
 ACCESSION AC018410.24 GI:24850147
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 147024)
 Birren, B., Nusbaum, C. and Lander, E.
 Homo sapiens chromosome 11, clone RP11-17G12
 Unpublished
 2 (bases 1 to 147024)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Beckert, R., Bedalov, F.,
 Boguslavskiy, L., Bouckgalter, B., Brown, A., Castle, A., Colangelo, M.,
 Collins, S., Collymore, A., Cooke, P., DeRubeis, K., Dewar, K.,
 Domino, M., Doyle, M., Fenesfor, J., Ferreira, P., Fitzhugh, W.,
 Forrest, C., Gage, D., Galagan, J., Gaidy, S., Grant, G., Hagos, B.,
 Heath, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L.,
 Karlas, A., Klein, J., Landers, T., Lebecky, J., Liu, C., Locke, K.,
 Macdonald, P., Margulis, N., McBryan, P., McElroy, A., McKernan, K.,
 Melartin, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T.,
 O'Donnell, P., Peterson, K., Pierre, N., Pollara, V., Riley, R.,
 Rothman, D., Roy, A., Santos, R., Severy, P., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, D., Testa, S., Theodore, J.,
 Tirrell, A., Vassiliou, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
 Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (10-Dec-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 147024)
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Bouckgalter, B.,

Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pterre, N., Raymond, C., Retta, R., Risse, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testfaye, S., Theodores, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (04-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 147024)

REFERENCE
JOURNAL
AUTHORS
Birn, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Chepel, Y., Collymore, A., Cooke, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pterre, N., Raymond, C., Retta, R., Risse, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testfaye, S., Theodores, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (19-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
5 (bases 1 to 147024)

REFERENCE
JOURNAL
AUTHORS
Birn, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Chepel, Y., Collymore, A., Cooke, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pterre, N., Raymond, C., Retta, R., Risse, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testfaye, S., Theodores, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (10-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 10, 2002 this sequence version replaced g1:24137549.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information

Center project name: L3636
Center clone name: 17_g_12

Only the last 147.0 kilobases of this clone are being submitted.
The remainder overlaps accession number AC090582 [WTCGR project L11783].

FEATURES
source
1. 147024
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-17G12"
/clone_1kb="RP11-11 Human Male BAC"
complement(1..60)
/rpt_family="Alu"
complement(100..193)
/rpt_family="L2"
345..372
/rpt_family="AT-rich"
426..732
/rpt_family="AluSg"
complement(733..999)
/rpt_family="AluO"
1033..1054
/rpt_family="AT-rich"
complement(1058..1286)
/rpt_family="MIR"
complement(12149..2459)
/rpt_family="AluY"
complement(2463..2753)
/rpt_family="AluSc"
complement(2786..2981)
/rpt_family="MIR"
2994..3151
/rpt_family="MIR"
complement(3975..4138)
/rpt_family="MER91A"
4557..4735
/rpt_family="(TG)n"
4766..4950
/rpt_family="(TG)n"
4996..5179
/rpt_family="(TG)n"
5201..5313
/rpt_family="(TG)n"
5339..5434
/rpt_family="(TG)n"
5435..5727
/rpt_family="AluSg"
5728..5857
/rpt_family="(TG)n"
6048..6201
/rpt_family="(TG)n"
6282..6458
/rpt_family="(TG)n"
6477..6603
/rpt_family="(TG)n"
6766..6942
/rpt_family="(TG)n"
7001..7180
/rpt_family="(TG)n"
7216..7288
/rpt_family="(TG)n"
7305..7480
/rpt_family="(TG)n"
7492..7611
/rpt_family="(TG)n"
8199..8256
/rpt_family="C-rich"
8724..8751
repeat_region

```
repeat_region      /rpt_family="ATG)n"
                    complement(8800..8932)
                    /rpt_family="MIR3"
repeat_region      /rpt_family="MIR3"
                    complement(9403..9710)
                    /rpt_family="Aluub"
repeat_region      /rpt_family="Aluub"
                    complement(9712..9788)
                    /rpt_family="AT_rich"
repeat_region      /rpt_family="Aluub"
                    complement(9789..10086)
                    /rpt_family="Aluub"
repeat_region      /rpt_family="MIR"
                    complement(10094..10269)
                    /rpt_family="MIR"
repeat_region      /rpt_family="MIR"
                    complement(10275..10340)
                    /rpt_family="crRNA-Met-1"
repeat_region      /rpt_family="MIR3"
                    complement(11030..11164)
                    /rpt_family="MIR3"
repeat_region      /rpt_family="MIR3"
                    complement(11536..11720)
                    /rpt_family="MIR"
repeat_region      /rpt_family="MIR"
                    complement(12144..12212)
                    /rpt_family="L2"
```

Query Match 7.5%; Score 374.2; DB 9; Length 147024;

Best Local Similarity 68.1%; Pred. No. 8.2e-61;

Matches 637; Conservative 0; Mismatches 263; Indels 35; Gaps 7;

```
QY 5 GGCACATGCTGTGGCTCAGCGGTATTCGCCACTTGGG-----AGGACGAGGAT 58
DB 79292 GGCACGAGACATGCGCCCTGTAATCAAGCACTTGGACGCTAAGCGGAT 79351
QY 59 CCCTTGAGCTCAGATTGAGACAGCCTAGTAAATATGTAACCTCTGCTGTACA 118
DB 79352 CACCTGAGGCTGTGAGTTGAGACAGCCTGGCAACATGTGAA-CAACATCTTACT 79410
QY 119 AATAATAAAGATTTTCCAGGATGTGGCTGCACCCCAAGTGCAGATTTGGAG 178
DB 79411 AAAAAATACAAAAAGTAGCCGGGTGAGTGGCTCGCCCTGTATCTTACACTTTGGAG 79470
QY 179 GGTGAGTGAAGGATGCTTGAAGCAGAGGATGTAAGCAAGCCTAGGCAACATGTGA 238
DB 79471 GCGAGGACGAGGAAATGCTTGAAGCATGAGTTGAAACAGCTGCGCAACATGTGA 79530
QY 239 GACCTGTGTCTATTA-----AATAATTAGCTGTGTCTTGGCAGAGCTGCA-- 290
DB 79531 AACCTGTGTCTACTATAAACAACAAAAATTAGCTGGGCTGATGCACTGTCTTAAT 79590
QY 291 GCTAGCTACTCGAAGACTGAGGTGGAGATCAC-TGAGCCGAGAGGCTGAGCTGCA 349
DB 79591 CCGAGCTACTCGGGTGGCTGGGCAAGAAATCGCTTGAACCGAGAGCGAGTTGCA 79650
QY 350 GTGAACAGTATCACCAGCTGATTCGAGCTGGAAGACAGAGGAGACCTGTTCCA 409
DB 79651 GTGAGCCGAGATCAGCAGCTGCTGAGCTGGGCGACAGAGCAAGATCTCATCTCA 79710
QY 410 AAAAAAAAAAAAAAAAAAATGCAAGAAAGACATCATTAACCTTGACTGGGACATTA 469
DB 79711 AAAAAATATATATACAAAAT-----TAGCAGGCAAGTGGCAGCAC 79754
QY 470 TTTATGTATGAATTCACATCTTTTAGAGAAATTAAGCATTTCTGATTAATGAT 529
DB 79755 TGTAGTCCAGCTACTTGGAGGCGGAGGAGAGAAATCATTTGAACCTGGAGGAGAA 79814
QY 530 AATAATTATTTTATTAATTCAAATGAAATTAATTAATCTGAGAACTAGCTTCACT 589
DB 79815 CTTCAGTATGCTGAGACGCGCACACTGACCTCCAGACCGGCTGACAGACAGCA 79874
QY 590 CTCTCAGTTGCTAGTCAAACTTTAATGCTTTTGGCGGGTGGCTGCTCAGCGCTGT 649
DB 79875 CTGTCACT--ACATACAAAAATTAATTAAGCCCGGCGGCTGCTCAGCGCTGT 79931
QY 650 AATCCAGCATTTGGGAGCGCGAGCGGGTGGATCAAGGTTAGAGATGAAGCAT 709
DB 79932 ATCTCAGCATTTGGGAGCGCGAGCGGGGAGATCAAGGTCAGAGATCGAGATCAT 79991
QY 710 CCTGGCTAACAGGTAACCTGTCTCTACTAAAAATACAAAAATTAAGCCGGGTGGG 769
```

```
DB 79992 CCTGGCTAACAGGTAACCTGTCTCTACTAAAAATACAAAAATTAAGCCGGGTGG 80051
QY 770 TCCAGAGCGCTGTAGTCCAGCTGCTCAGAGGCTGAGGACAGAGATGTGTGAACCC 829
DB 80052 TGGCGGGCGCTGTAGTCCAGCTACTCGGAGGCTGAGGACAGAGATGGCATGAAGCC 80111
QY 830 GGGAGCGGAGCTTGCAGTGAAGCCGAGATTTGGCCACTGCACCTCGGCGCAGAT 889
DB 80112 GGGAGCGGAGCTTGCAGTGAAGCCGAGATTTGCTCACTGCACCTCGGCGTACAGA 80171
QY 890 GCGAGCTGTCTCTCAAAAAAAAAAAAAAAAAA 924
DB 80172 GCGAGCTGTCTCTCAAAAAAAAAATTAATTA 80206
```

Search completed: December 27, 2004, 04:17:57
Job time : 13663.4 secs

This Page Blank (uspto)